LONG TERM DYNAMICS OF STOCHASTIC EVOLUTION EQUATIONS

Proefschrift

ter verkrijging van

de graad van Doctor aan de Universiteit Leiden,

op gezag van de Rector Magnificus prof. mr. P.F. van der Heijden,

volgens besluit van het College voor Promoties

te verdedigen op dinsdag 9 februari 2010

klokke 16.15 uur

door

Gregorius Nicolaas Johannes Cornelis Bierkens

geboren te Tilburg

in 1979

Samenstelling van de promotiecommissie:

Promotor:	prof. dr. S.M. Verduyn Lunel
Copromotor:	dr. ir. O. van Gaans
Overige leden:	prof. dr. ir. J.H. van Schuppen (Technische Universiteit Delft)
	prof. dr. A. Chojnowska-Michalik (Uniwersytet Łódzki, Polen)
	prof. dr. W.Th.F. den Hollander
	prof. dr. P. Stevenhagen

Long Term Dynamics of
Stochastic Evolution Equations

Bierkens, Joris, 1979 –
Long Term Dynamics of Stochastic Evolution Equations
Printed by Ipskamp Drukkers, Enschede
Cover illustration by Vesta Kroese (lino print)
AMS 2010 Subj. class code: 60H99, 34K50, 60H15
NUR: 918
ISBN: 978-90-9025010-6

email: joris@math.leidenuniv.nl

THOMAS STIELTJES INSTITUTE
FOR MATHEMATICS

CONTENTS

CHAPTER

ONE

Introduction

Stochastic differential equations with delay are the inspiration for this thesis. Examples of such equations arise in population models, control systems with delay and noise, lasers, economical models, neural networks, environmental pollution and in many other situations. In such models we are often interested in the *evolution* of a particular quantity, for example the size of a population, or the amount of pollution in a particular area, changing in time.

A differential equation with delay, or *delay equation*, is a differential equation in which the change in time of such a quantity is expressed as a function of the value of that quantity at different points in time, in the past as well as in the present. This is in contrast with an ordinary differential equation, in which the change in time of the quantity at a specific time is expressed as a function of that quantity at that specific time only.

We may add the influence of uncertainty or noise to such a delay equation. We do this by making the change in time of the quantity also dependent on a noise process such as a Wiener process or a Poisson process. The differential equation obtained in this way is called a stochastic differential equation with delay or *stochastic delay differential equation*. The change in time of some quantity under uncertainty is called a *stochastic evolution*. Henceforth, when we speak of a differential equation, the reader should keep this type of equation in mind.

Just describing a model, or, in mathematical language, composing a system of differential equations, is little satisfactory. We want to be able to say something about properties of the model, or equivalently the qualitative behaviour of the solutions of the system of differential equations.

Ideally we would be able to determine the explicit solution of a given differential equation. This is the complete description of future behaviour of the evolution, as a function of time and the sources of uncertainty. However this is almost never possible for the differential equations which are the subject of this thesis.

Instead, to obtain a better understanding of the qualitative behaviour of solutions, we can study *stationary behaviour*. Stationary behaviour is behaviour which, once it arises, will repeat itself. Consider for example a constant value or a periodic solution.

To understand behaviour we first need to understand what possible *states* the solution of a stochastic delay differential equation can attain. Or equivalently, what is the *state space* of such an equation? Because for delay equations the change in time of the quantity depends on the past, in order to know the future we must know the past. Therefore the *state* of the quantity does not only consist of its present value, but also the values in the past which we require for the description of the future of the quantity. As a consequence the state space is a function space, which is an infinite dimensional space. Therefore also stochastic delay differential equations have an infinite dimensional state space.

By now a large amount of theory exists on the subject of infinite dimensional stochastic differential equations. However, this theory is often not applicable to stochastic delay equations due to their specific peculiarities. Indeed, the semigroup corresponding to a delay equation is only eventually compact (and not immediately, as in the case of the heat semigroup). Also the noise of stochastic delay differential equations is degenerate: the noise only influences the stochastic evolution in certain directions, not all, which is often the case with stochastic partial differential equations. This degeneracy occurs because the noise cannot influence the past, which is however part of our state space.

After this introduction we first describe stochastic differential equations in infinite dimensions (Chapter 2). In the subsequent chapter (Chapter 3) we use this theory for the definition of stochastic delay differential equations. As an ingredient we will introduce the *delay semigroup*, which describes solutions of linear delay equations. In that chapter we also describe some properties of the delay semigroup: it being eventually compact, and the fact that the inner product on the state space can be chosen in such a way that the delay semigroup is a generalized contraction.

We may then embark on the study of long term dynamics of stochastic delay equations. For this we return to the notion of stationary behaviour. For stochastic differential equations, stationary behaviour is characterized by a probability distribution on the state space which is invariant under the stochastic evolution: once the solutions of the stochastic differential equation have this probability distribution the future values of the solution have that same particular probability distribution. Such a probability distribution is called an *invariant probability measure*.

In Chapter 4 it is shown that, using the eventual compactness of the delay semigroup, the existence of an invariant probability measure can be shown under reasonable conditions.

In Chapter 5 conditions are given under which the invariant probability measure is unique, using techniques from Malliavin calculus. If there exists only one invariant probability measure then, due to the ergodic principle, the average long term behaviour will be given by this invariant probability measure.

Finally, in Chapter 6, the long term behaviour of linear stochastic evolutions with multiplicative noise is investigated. In particular conditions for the (pathwise) stability of solutions are given; first for the case of non-degenerate noise. Then we study the case relevant for stochastic delay equations, namely the case of degenerate noise.

A population model with random migration

We finish this introductory chapter with an example to illustrate part of the theory developed in this thesis.

Consider a simple model for the evolution of the size of a population of animals. Each year the amount of newly born individuals is equal to a fraction β (the *birth rate*) of the size of the adult population. Also each year a fraction α (the *death rate*) of the population dies.

For simplicity we assume that it takes one year for an individual to reach maturity. If we assume that only a small number of deaths occur in the time between birth and maturity, then the size of the adult population at time t is roughly equal to the size of the total population at time $t-1$.

If the birth rate exceeds the death rate, then the population will grow in time, see Figure 1.1(a) for an example.

Suppose now that random migration with a neighbouring population takes place: every year un unknown amount of individuals migrate between the two populations. The immigration is unbiased: the expected net migration between the two populations is zero. Furthermore the migration is proportional to the size of the population with propotionality constant σ.

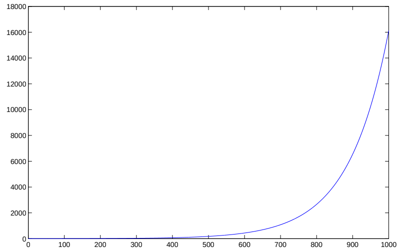
We will not go into further details regarding the exact nature of the random effects here. However, for the interested reader, we give the *stochastic delay differential equation* governing this evolution:

$$dx(t) = [-\alpha x(t) + \beta x(t-1)]\ dt + \sigma x(t)\ dW(t) \tag{1.1}$$

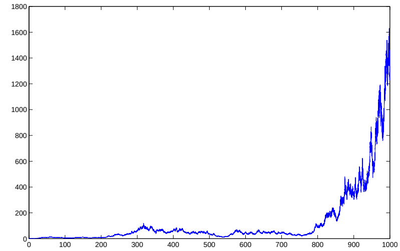
with W a standard Brownian motion. A precise description is given in Section 3.3.1. The necessary background material for this is described in Chapter 2 and Chapter 3 where stochastic differential equations and stochastic differential equations with delay are introduced.

Stability

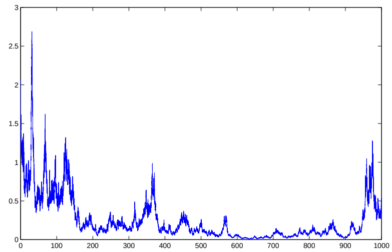
The following interesting phenomenon occurs: For relatively large rates of random migration the size of the population tends to zero with probability one, even if the birth rate exceeds the death rate. See Figure 1.1(d). We say that in this case the stochastic evolution is *stable*.



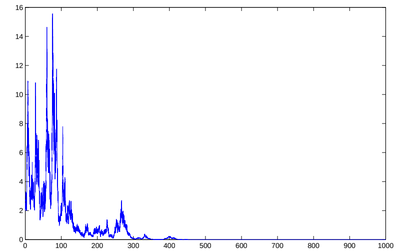
(a) No migration, $\sigma = 0$



(b) Random migration at rate $\sigma = 0.1$



(c) Random migration at rate $\sigma = 0.2$



(d) Random migration at rate $\sigma = 0.3$

Figure 1.1: The evolution of the size of a population in time with birth rate $\beta = 0.11$ and death rate $\alpha = 0.1$. Without migration or for a small rate of random migration the population increases steadily in size, whereas for a larger rate of migration the population is eventually extinguished.

These results are in line with theory developed in Chapter 6; see Example 6.30.

CHAPTER

TWO

Stochastic differential equations

This thesis deals with stochastic evolutions. More specifically, we study solutions of stochastic differential equations, for which we need the notion of the stochastic integral.

The goal of this chapter is to introduce the concepts 'stochastic integral' and 'stochastic differential equation' in a manner that is general enough to allow us to study stochastic evolutions driven by many Lévy processes (square integrable Lévy processes) and by cylindrical Wiener processes. In Section 2.1 we introduce the notion of square integrable martingales. The stochastic integral with respect to these processes is defined in Section 2.2. To be able to study equations driven by more general noise processes we introduce the stochastic integral with respect to cylindrical martingales in Section 2.3. Some examples of these processes are given in Section 2.4. In Section 2.5 we describe what we mean by a solution of a stochastic differential equation and mention some regularity results.

Throughout this thesis, let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ denote a filtered probability space satisfying the usual conditions (of completeness and right continuity, see [73]), and let $(U, \langle\cdot,\cdot\rangle_U, |\cdot|_U)$ and $(H, \langle\cdot,\cdot\rangle_H, |\cdot|_H)$ be Hilbert spaces. We will assume that U is separable (but H not necessarily). When this does not lead to confusion we will omit the subscripts in $\langle\cdot,\cdot\rangle$ and $|\cdot|$.

2.1 Square integrable martingales

A U-valued stochastic process $X = (X(t))_{t \geq 0}$ is called *integrable* if $\mathbb{E}|X(t)| < \infty$ for all $t \geq 0$. If X is integrable, adapted and if

$$\mathbb{E}\left[X(t)|\mathcal{F}_s\right] = X(s), \quad \mathbb{P}\text{-a.s., } 0 \leq s \leq t,$$

then X is called a *martingale*. If $\mathbb{E}\left[|X(t)|^2\right] < \infty$ for $t \geq 0$ then X is called *square integrable*. A stochastic process process X is called *cadlag* (continu à droite et pourvu de limites à gauche) if it is *right-continuous* and has left limits almost surely, that is,

$$X(t+) := \lim_{s \downarrow t} X(s) = X(t) \quad \text{and} \quad X(t-) := \lim_{s \uparrow t} X(s) \text{ exists almost surely.}$$

We denote the space of all cadlag square integrable martingales in U adapted to $(\mathcal{F}_t)$ by $\mathcal{M}^2(U)$.

Let $M \in \mathcal{M}^2(U)$. Then $|M|_U^2$ is a submartingale in $\mathbb{R}$, and hence by the Doob-Meyer decomposition there exists a unique increasing, predictable real-valued process $\langle M \rangle$ such that $|M|_U^2 - \langle M \rangle$ is a martingale. Using polarization define the predictable real-valued process $\langle M, N \rangle := \frac{1}{4}\left(\langle M + N \rangle - \langle M - N \rangle\right)$ for $M, N \in \mathcal{M}^2(U)$. It can then be verified that $\langle M, N \rangle_U - \langle M, N \rangle$ is a martingale.

Recall the notion of a nuclear operator (see Appendix A). Let $L_1(U)$ denote the separable Banach space of nuclear operators on U equipped with the corresponding norm and let $L_1^+(U)$ denote the cone in $L_1(U)$ consisting of all self-adjoint non-negative nuclear operators. For $x, y, z \in U$ we denote the mapping $z \mapsto \langle y, z \rangle_U x$ by $x \otimes y$. We have $||x \otimes y||_{L_1(U)} = |x|_U |y|_U$.

If $M \in \mathcal{M}^2(U)$ then the process $M \otimes M$ is an $L_1(U)$-valued right-continuous process such that

$$\mathbb{E}||M(t) \otimes M(t)||_{L^1(U)} = \mathbb{E}|M(t)|_U^2 \leq \mathbb{E}|M(T)|_U^2 < \infty, \quad 0 \leq t \leq T.$$

A process X in $L(U)$ is said to be *increasing* if $X(t)$ is self-adjoint and nonnegative, $t \geq 0$, and $X(s) \leq X(t)$ almost surely for $0 \leq s \leq t$ with respect to the partial order induced by the cone of nonnegative operators.

We mention the following result ([62], Théorème 1, or [71], Theorem 8.2).

Theorem 2.1. *Let $M \in \mathcal{M}^2(U)$. Then there is a unique right-continuous $L_1^+(U)$-valued increasing predictable process $\langle\langle M \rangle\rangle$ such that $\langle\langle M \rangle\rangle(0) = 0$ and the process $M \otimes M - \langle\langle M \rangle\rangle$ is an $L_1(U)$-valued martingale. Moreover there exists a predictable $L_1^+(U)$-valued process Q such that*

$$\langle\langle M \rangle\rangle(t) = \int_0^t Q(s)\ d\langle M \rangle(s). \tag{2.1}$$

The $L_1^+(U)$-valued process $\langle\langle M \rangle\rangle$ is called the *operator angle bracket* corresponding to M or the *quadratic variation* of M and the $L_1^+(U)$-valued process Q satisfying (2.1) is called the *martingale covariance process* of M. If $\langle\langle M \rangle\rangle(t) = Rt$ for some fixed $R \in L_1^+(U)$, then M is said to be of *stationary covariance* R.

In order to study regularity of paths of solutions of stochastic differential equations (see Section 2.5), we mention the following version of the well known *Burkholder-Davis-Gundy inequalities*:

Theorem 2.2 (Burkholder-Davis-Gundy inequalities). *Let U be a separable Hilbert space. For $0 < p < \infty$ there exist constants $0 < c_p \leq C_p$ such that for every continuous square integrable martingale M in U we have*

$$c_p \mathbb{E}\langle M \rangle^{p/2}(T) \leq \mathbb{E} \sup_{0 \leq t \leq T} |M(t)|_U^p \leq C_p \mathbb{E}\langle M \rangle^{p/2}(T).$$

Remark 2.3. See [75], Theorem IV.42.1, where the theorem is proven for the real-valued case. This can be extended to the Hilbert space valued case without trouble.

We will also make use of the following regularity result (see [71], Theorem 3.41):

Proposition 2.4. *Let M be a stochastically continuous square integrable martingale taking values in U. Then M has a cadlag modification satisfying*

$$\mathbb{P}\left(\sup_{t \in [0,T]} |M(t)|_U \geq r \right) \leq \frac{\mathbb{E}|M(T)|_U^2}{r^2}, \quad \textit{for all } T \geq 0, r > 0.$$

Moreover

$$\mathbb{E} \sup_{t \in [0,T]} |M(t)|_U^\alpha \leq \frac{2}{2-\alpha} (\mathbb{E}|M(T)|_U^2)^{\alpha/2}, \quad \textit{for all } T \geq 0, \alpha \in (0,2).$$

2.1.1 Wiener processes

A stochastic process X taking values in U has *independent increments* if, for any $0 \leq t_0 < t_1 < \ldots < t_n$ the U-valued random variables $X(t_1) - X(t_0), X(t_2) - X(t_1), \ldots, X(t_n) - X(t_{n-1})$ are independent.

Let $R \in L_1^+(U)$ and assume $R \neq 0$ to exclude the trivial case. A U-valued process W is called an *R-Wiener process*, or just *Wiener process* if

(i) $W(0) = 0$,

(ii) W has continuous trajectories,

(iii) W has independent increments,

(iv) $\mathcal{L}(W(t) - W(s)) = \mathcal{N}(0, (t-s)R)$, $0 \leq s \leq t$.

Here $\mathcal{L}(X)$ denotes the probability law of a random variable X, and $\mathcal{N}(m, C)$ denotes the normal distribution with mean $m \in U$ and covariance $C \in L_1^+(U)$. For arbitrary $R \in L_1^+(U)$ such a process exists ([24], Proposition 4.2). It is easily checked that $W \in \mathcal{M}^2(U)$ with quadratic variation process $\langle\langle W \rangle\rangle(t) = Rt$. Furthermore $\mathbb{E}|W(t)|_U^2 = \text{tr } R$ so the martingale covariance process of W (defined in Theorem 2.1) is $Q(t) = R/\text{tr } R$.

2.1.2 Lévy processes

Let L be a stochastic process in U with independent increments. If $\mathcal{L}(L(t) - L(s))$ depends only on the difference $t - s$ then we say that L has *stationary, independent increments*. If in addition $L(0) = 0$ and L is continuous in probability then L is called a *Lévy process*. By [71], Theorem 4.3, we can (and will) always choose a version of L which is cadlag.

Let ν be a finite Borel measure on a Hilbert space U such that $\nu(\{0\}) = 0$. A *compound Poisson process* with *Lévy measure* or *jump intensity measure* ν is a cadlag Lévy process L satisfying

$$\mathbb{P}(L(t) \in \Gamma) = \mathrm{e}^{-\nu(U)t} \sum_{k=0}^{\infty} \frac{t^k}{k!} \nu^{*k}(\Gamma), \quad t \geq 0, \Gamma \in \mathcal{B}(U).$$

Here the *convolution of measures* μ and ν is the Borel measure on U defined by

$$(\nu * \mu)(\Gamma) = \int_U \nu(\Gamma - x)\mu(dx), \quad \Gamma \in \mathcal{B}(U),$$

and ν^{*k} denotes the k-th power of convolution of measures, with $\nu^{*0} = \delta_0$, the Dirac measure at 0. For any finite Borel measure ν satisfying $\nu(\{0\}) = 0$ such a compound process can be constructed ([71], Theorem 4.15).

If L is an integrable compound Poisson process such that $\mathbb{E}L(t) = 0$ for all $t \geq 0$, then L is called a *compensated compound Poisson process*. To any integrable compound Poisson process L we can associate such a compensated compound Poisson process $\widehat{L}$ by defining $\widehat{L}(t) = L(t) - \mathbb{E}L(t)$, $t \geq 0$.

The following decomposition result is classical ([71], Theorem 4.23):

Theorem 2.5. *Let L be a cadlag Lévy process in U.*

(i) If ν is a jump intensity measure corresponding to a Lévy process then

$$\int_U \left(|y|_U^2 \wedge 1\right) \ \nu(dy) < \infty.$$

(ii) Every Lévy process has the following representation:

$$L(t) = at + W(t) + \sum_{n=1}^{\infty} L_n(t) + L_0(t), \quad t \geq 0, \tag{2.2}$$

where $a \in U$, W is a Wiener process in U, L_0 is a compound Poisson process with jump intensity measure $\mathbb{1}_{|y|_U \geq r_0}(y)\nu(dy)$ and each L_n is a compensated compound Poisson process with jump intensity measure $\mathbb{1}_{r_{n+1} \leq |y|_U < r_n}(y)\nu(dy)$. Here (r_n) is an arbitrary positive sequence decreasing to 0. Furthermore all members of the representation are independent processes and the series converges $\mathbb{P}$-a.s. uniformly on each bounded subinterval of $[0,\infty)$.

A cadlag Lévy process L is square integrable if and only if its Lévy measure satisfies

$$\int_U |y|_U^2 \ \nu(dy) < \infty.$$

Assume L is a square integrable Lévy process and L has representation (2.2). Then the *compensated Lévy process* $\widehat{L} := L - \mathbb{E}L$ is a square integrable martingale, satisfying the following properties:

$$\begin{aligned}
\mathbb{E}L(t) &= \left(a + \int_{\{|y|_U \geq r_0\}} y \ \nu(dy)\right) t, && t \geq 0,\\
\langle \widehat{L} \rangle(t) &= \mathbb{E}|\widehat{L}(t)|_U^2 = t \ \mathrm{tr}\ Q && t \geq 0,\\
\langle\langle \widehat{L} \rangle\rangle(t) &= \mathbb{E}\widehat{L}(t) \otimes \widehat{L}(t) = tQ, && t \geq 0.
\end{aligned}$$

where $Q = Q_0 + Q_1$ with Q_0 the covariance operator of the Wiener part W of L and Q_1 is given by

$$\langle Q_1 x, z\rangle_U = \int_U \langle x, y\rangle_U \langle z, y\rangle_U \ \nu(dy), \quad x, z \in U.$$

See [71], Theorem 4.47.

Remark 2.6. Lévy process are rarely square integrable. For example, the only example of a square integrable Lévy process in the important class of stable processes is Brownian motion.

2.2 Stochastic integral with respect to square integrable martingales

An $L(U;H)$-valued process Ψ is said to be *simple* if there exists a sequence of non-negative reals $t_0 = 0 < t_1 < \ldots < t_m$, a sequence of operators $\Psi_i \in L(U;H)$, $i = 0, \ldots, m-1$, and a sequence of events $A_i \in \mathcal{F}_{t_i}$, $i = 0, \ldots, m-1$, such that

$$\Psi(t) = \sum_{i=0}^{m-1} \mathbb{1}_{A_i} \mathbb{1}_{(t_i, t_i+1]}(t) \Psi_i, \quad t \geq 0.$$

Let $\mathcal{S}(U;H)$ denote the set of simple processes with values in $L(U;H)$. For a simple process Ψ define the stochastic integral of Ψ with respect to $M \in \mathcal{M}^2(U)$ by

$$\int_0^t \Psi \, dM := \sum_{i=0}^{m-1} \mathbb{1}_{A_i} \Psi_i \left(M(t_{i+1} \wedge t) - M(t_i \wedge t) \right), \quad t \geq 0.$$

Let $L_{\mathrm{HS}}(U;H)$ denote the space of all Hilbert-Schmidt operators from U into H equipped with the corresponding norm (see Appendix A, and note in particular that H need not be separable. Then ([62], Théorème 2; [71], Proposition 8.6) the following form of the *Itô isometry* holds:

$$\mathbb{E} \left| \int_0^t \Psi \, dM \right|_H^2 = \mathbb{E} \int_0^t \left|\left| \Psi \circ Q^{1/2} \right|\right|_{L_{\mathrm{HS}}(U;H)}^2 \, d\langle M \rangle. \tag{2.3}$$

For $T < \infty$, equip $\mathcal{S}(U;H)$ with the seminorm

$$||\Psi||_{M,T}^2 := \mathbb{E} \int_0^T ||\Psi \circ Q^{1/2}||_{L_{\mathrm{HS}}(U;H)} \, d\langle M \rangle. \tag{2.4}$$

Now let $L_{M,T}^2(H)$ be the completion of $(\mathcal{S}, ||\cdot||_{M,T}^2)$ with respect to the norm defined by (2.4), and as usual identify $\Psi, \Phi \in L_{M,T}^2(H)$ whenever $||\Psi - \Phi||_{M,T} = 0$.

We will characterize the space $L_{M,T}^2(H)$ for the case in which M is of stationary covariance.

Let $\mathcal{P}_T$ denote the σ-algebra of *predictable sets* , that is, the σ-algebra of subsets of $[0,T] \times \Omega$ generated by sets of the form $(s,t] \times A$, where $0 \leq s < t \leq T$ and $A \in \mathcal{F}_s$.

Theorem 2.7. *Suppose $M \in \mathcal{M}^2(U)$ is of stationary covariance Q. Then*

$$L_{M,T}^2(H) = \left\{ \Phi \circ Q^{-1/2} : \Phi \in L^2(\Omega \times [0,T], \mathcal{P}_T, d\mathbb{P} \otimes dt; L_{\mathrm{HS}}(U;H)) \right\}, \tag{2.5}$$

and, for $\Psi = \Phi \circ Q^{-1/2}$ with Φ as in the characterization above,

$$||\Psi||_{M,T}^2 = \mathbb{E} \int_0^T ||\Phi||_{L_{\mathrm{HS}}(U;H)}^2 \, dt. \tag{2.6}$$

Proof Let $\mathcal{V}$ denote the righthandside of (2.5). If $\Psi \in \mathcal{S}(U;H)$, then we may define $\Phi := \Psi \circ Q^{1/2}$ and it is immediately clear that $\Phi \in L^2(\Omega \times [0,T], \mathcal{P}_T, d\mathbb{P} \otimes dt; L_{\mathrm{HS}}(U;H))$. So $\Psi \in \mathcal{V}$. It is straightforward to check that (2.6) holds on $\mathcal{S}(U;H)$.

Now assume $\Phi \in L^2(\Omega \times [0,T], \mathcal{P}_T, d\mathbb{P} \otimes dt; L_{\mathrm{HS}}(U;H))$. Then there exists a sequence (Φ_n) of simple, predictable functions such that $\Phi_n \to \Phi$ in $L^2(\Omega \times [0,T]; L_{\mathrm{HS}}(U;H))$ by the definition of Bochner integrable functions. Let (e_i) be a complete orthonormal system of eigenvectors of Q, and let Π_n denote the orthogonal projection on the finite dimensional subspace spanned by $\{e_1, \ldots, e_n\}$. Then $\Phi_n \Pi_n \to \Phi$ in $L^2(\Omega \times$

$[0,T]; L_{\mathrm{HS}}(U;H))$, and $\Phi_n \Pi_n Q^{-1/2} \in \mathcal{S}(U;H)$ for almost all $t \in [0,T]$. This shows that $\mathcal{S}(U;H)$ is dense in $\mathcal{V}$. (Note that $\Phi_n Q^{-1/2}$ does not necessarily assume values in $L(U;H)$, however $\Phi_n \Pi_n Q^{-1/2}$ does.) □

Remark 2.8. A similar result holds for general $M \in \mathcal{M}^2(U)$, see [71], Theorem 8.10.

We can now define the stochastic integral $\Psi \mapsto \int_0^{\cdot} \Psi \, dM$ for processes in $L^2_{M,T}(H)$ as follows ([62], Théorème 2 or [71], Theorem 8.7).

Theorem 2.9. *(i) There exists a unique extension of the mapping $\mathcal{S}(U;H) \to L^2(\Omega, \mathcal{F}_T, \mathbb{P}; H)$, $\Psi \mapsto \int_0^T \Psi \, dM$ to an isometry from $L^2_{M,T}(H)$ into $L^2(\Omega, \mathcal{F}_T, \mathbb{P}; H)$, also denoted by $\Psi \mapsto \int_0^T \Psi \, dM$ and called the* stochastic integral *of Ψ with respect to M.*

(ii) For all $\Psi \in L^2_{M,T}(H)$ and $0 \le s \le t \le T$ we have $\mathbb{1}_{(s,t]}\Psi \in L^2_{M,t}(H)$ and

$$\mathbb{E}\left|\int_0^t \Psi \, dM - \int_0^s \Psi \, dM\right|_H^2 = ||\mathbb{1}_{(s,t]}\Psi||^2_{M,t} \le ||\Psi||^2_{M,T}.$$

We write

$$\int_s^t \Psi \, dM := \int_0^t \Psi \, dM - \int_0^s \Psi \, dM, \quad 0 \le s \le t \le T.$$

(iii) For any $\Psi \in L^2_{M,T}(H)$ the process $\left(\int_0^t \Psi \, dM\right)_{t \ge 0}$ is an H-valued square integrable, mean-square continuous martingale starting from 0.

(iv) For any $\Psi, \Phi \in L^2_{M,T}(H)$ and any $t \in [0,T]$,

$$\left\langle \int_0^{\cdot} \Psi \, dM, \int_0^{\cdot} \Phi \, dM \right\rangle (t) = \int_0^t \langle \Psi \circ Q^{1/2}, \Phi \circ Q^{1/2} \rangle_{L_{\mathrm{HS}}(U;H)} \, d\langle M \rangle$$

and

$$\left\langle\!\!\left\langle \int_0^{\cdot} \Psi \, dM, \int_0^{\cdot} \Phi \, dM \right\rangle\!\!\right\rangle (t) = \int_0^t \Psi \circ Q \circ \Phi^* \, d\langle M \rangle.$$

Remark 2.10. In the case where M is a square integrable martingale of stationary covariance Q, we have by Theorem 2.9 and equation (2.6) the following version of the Itô isometry:

$$\mathbb{E}\left|\int_0^T \Psi \, dM\right|^2 = \mathbb{E}\int_0^T ||\Psi \circ Q^{1/2}||^2_{L_{\mathrm{HS}}(U;H)} \, dt. \tag{2.7}$$

Proposition 2.11. *Suppose A is a closed linear operator from $\mathfrak{D}(A) \subset H$ into H. Suppose that $\mathfrak{D}(A)$ is a Borel subset of H. Let $\Psi \in L^2_{M,T}(H)$, and suppose that for almost all $t \in [0,T]$ we have that $\Psi(t) \circ Q(t)^{1/2}$ maps into $\mathfrak{D}(A)$ almost surely. Furthermore assume that*

$$||A\Psi||^2_{M,T} < \infty.$$

Then for $0 \le t \le T$ we have that $\int_0^t \Psi \, dM \in \mathfrak{D}(A)$ almost surely, and

$$A \int_0^t \Psi \, dM = \int_0^t A\Psi \, dM, \quad \textit{almost surely.}$$

PROOF: Recall that $\mathfrak{D}(A)$ equipped with the graph inner product, defined by

$$\langle x, y \rangle_{\mathfrak{D}(A)} := \langle x, y \rangle_H + \langle Ax, Ay \rangle_H, \quad x, y \in \mathfrak{D}(A),$$

is a Hilbert space.

Note that $\Psi \in L^2_{M,T}(\mathfrak{D}(A))$, since, for $T \in L(U; H)$, we have $T \in L_{\mathrm{HS}}(U; H)$ and $AT \in L_{\mathrm{HS}}(U; H)$ if and only if $T \in L_{\mathrm{HS}}(U; \mathfrak{D}(A))$ and

$$||T||^2_{L_{\mathrm{HS}}(U;H)} + ||AT||^2_{L_{\mathrm{HS}}(U;H)} = ||T||^2_{L_{\mathrm{HS}}(U;\mathfrak{D}(A))}.$$

We can now approximate Ψ by a sequence $(\Psi_n)_{n\in\mathbb{N}} \subset \mathcal{S}(U; \mathfrak{D}(A))$ in the $L^2_{M,T}(\mathfrak{D}(A))$-norm. For all $n \in \mathbb{N}$ we have

$$A \int_0^t \Psi_n \, dM = \int_0^t A\Psi_n \, dM, \quad \text{almost surely for all } 0 \le t \le T,$$

and

$$\int_0^t \Psi_n \, dM \to \int_0^t \Psi \, dM \quad \text{and} \quad \int_0^t A\Psi_n \, dM \to \int_0^t A\Psi \, dM$$

in $L^2(\Omega; \mathfrak{D}(A))$ and $L^2(\Omega; H)$, respectively, as $n \to \infty$. □

Proposition 2.12. *Suppose $M \in \mathcal{M}^2(U)$ has a continuous modification. Then for $\Psi \in L^2_{M,T}(H)$ the integral process $\left(\int_0^t \Psi \, dM\right)_{t\ge 0}$ has a continuous modification.*

PROOF: Let $(\Psi_n)_{n\in\mathbb{N}}$ a sequence in $\mathcal{S}(U; H)$ such that $\Psi_n \to \Psi$ in $L^2_{M,T}(H)$. Then, by Proposition 2.4, for $T > 0$ and $n \to \infty$,

$$\begin{aligned}\mathbb{E} \sup_{0\le t\le T} \left| \int_0^t \Psi_n - \Psi \, dM \right| &\le 2 \left(\mathbb{E} \left| \int_0^T \Psi_n - \Psi \, dM \right|^2_H \right)^{1/2} \\ &= 2 \left(\mathbb{E} \int_0^T ||(\Psi_n - \Psi) \circ Q||^2 \, d\langle M \rangle \right)^{1/2} \to 0.\end{aligned}$$

Hence for a subsequence (n_k),

$$\lim_{k\to\infty} \sup_{0\le t\le T} \left| \int_0^t \Psi_{n_k} - \Psi \, dM \right| = 0, \quad \text{almost surely.}$$

Hence $\int_0^\cdot \Psi \, dM$ is a uniform limit of continuous functions, almost surely. □

2.3 Cylindrical martingales

The theory of cylindrical martingales developed below is classical; a clear exposition can be found in [61].

Let V be a separable Hilbert space. A *cylindrical process* X is a family $(X(t))_{t\geq 0}$ such that for all $t \geq 0$, we have $X(t) \in L(V; L^2(\Omega, \mathcal{F}_t, \mathbb{P}))$. A cylindrical process X is called a *cylindrical (weakly cadlag) martingale* if $Xv = (X(t)v)_{t\geq 0}$ is a (cadlag) martingale for all $v \in V$. Those cylindrical martingales M for which, for some self-adjoint non-negative $R \in L(V)$, $(M(t)x)(M(t)y) - t\langle Rx, y\rangle$ is a martingale for all $x, y \in V$, are said to be of *stationary covariance* R. To avoid technical complications we will assume throughout that R is injective.

Example: Cylindrical Wiener process

A *cylindrical Wiener process* is a cylindrical martingale W defined by

$$(Wh)(t) = \sum_{j=1}^{\infty} \langle h, e_j\rangle_V W_j(t), \quad t \geq 0, h \in V,$$

where (e_j) is a complete orthonormal system for V, (W_j) are independent one-dimensional Wiener processes of variance $\sigma_j > 0$, for $j \in \mathbb{N}$, with $\sup_{j\in\mathbb{N}} |\sigma_j| < \infty$. Then W is a continuous cylindrical Lévy process of stationary covariance R, given by

$$\langle Rv, w\rangle = \sum_{j,k=1}^{\infty} \sigma_j \sigma_k \langle v, e_j\rangle_V \langle w, e_k\rangle_V, \quad v, w \in V.$$

Cylindrical Wiener processes occur as noise processes in stochastic partial differential equations, see Example 2.4.1. ◇

Suppose M is a cylindrical martingale of stationary covariance. We associate to M a square integrable martingale in some Hilbert space U, using the following variant of [71], Proposition 7.9.

Proposition 2.13. *Let U be a separable Hilbert space such that there exists a linear injective Hilbert-Schmidt mapping $J \in L_{\mathrm{HS}}(V; U)$, and note that $J^* : U^* \to V$. Suppose $Z \in L(V; L^2(\Omega, \mathcal{F}, \mathbb{P}))$ such that $\mathbb{E}(Zx)^2 = \langle Rx, x\rangle_V$ for some $R \in L(V)$, and $\mathbb{E}Zx = 0$ for $x \in V$. Then there is a unique mean-zero random variable $\widetilde{Z} \in L^2(\Omega, \mathcal{F}, \mathbb{P}; U)$ with covariance operator $JRJ^* \in L_1(U)$ such that*

$$ZJ^*\varphi = \varphi(\widetilde{Z}) \quad \text{for } \varphi \in U^*. \tag{2.8}$$

Proof: Let (e_j) be a complete orthonormal system in V. Note that $J^* \in L_{\mathrm{HS}}(U; V)$ by Remark A.5, (ii). Moreover, $JR \in L_{\mathrm{HS}}(U; V)$ by Proposition A.3, and $JRJ^* \in L_1(U)$ and tr $JRJ^* =$ tr J^*JR by Proposition A.4.

Define the random variable $\widetilde{Z}$ by

$$\widetilde{Z} = \sum_{j=1}^{\infty} (Ze_j) Je_j.$$

Then

$$\begin{aligned}
\mathbb{E}|\widetilde{Z}|_U^2 &= \sum_{j,k=1}^{\infty} \mathbb{E}(Ze_j)(Ze_k)\langle Je_j, Je_k\rangle_U = \sum_{j,k=1}^{\infty} \langle Re_j, e_k\rangle_V \langle Je_j, Je_k\rangle_U \\
&= \sum_{j=1}^{\infty} \left\langle Je_j, \sum_{k=1}^{\infty} \langle Re_j, e_k\rangle_V Je_k \right\rangle_U = \sum_{j=1}^{\infty} \langle Je_j, JRe_j\rangle_U = \mathrm{tr}\ J^* JR \\
&= \mathrm{tr}\ JRJ^*.
\end{aligned}$$

Expression (2.8) follows immediately and hence $\widetilde{Z}$ is unique. □

It is always possible to find a mapping J as mentioned in the previous proposition.

Proposition 2.14. *Let V be a separable Hilbert space. There exists a separable Hilbert space U and an injective mapping $J \in L_{\mathrm{HS}}(V;U)$.*

PROOF: Let (e_i) be a complete orthonormal system on V. Let $(\lambda_i) \in \ell^1(\mathbb{N})$ such that $0 < \lambda_i \leq 1$ for all $i \in \mathbb{N}$. For example, we may choose $\lambda_i = \frac{1}{i^2}$, $i \in \mathbb{N}$. Define a new inner product $\langle \cdot, \cdot \rangle_U$ on V by letting

$$\langle x, y\rangle_U = \sum_{i=1}^{\infty} \lambda_i \langle x, e_i\rangle_V \langle y, e_i\rangle_V.$$

Let U be the completion of V with respect to this inner product. Note that by the conditions on (λ_i), the inclusion $J : V \to U$ is bounded and injective. Furthermore,

$$||J||^2_{L_{\mathrm{HS}}(V;U)} = \sum_{i=1}^{\infty} |e_i|_U^2 = \sum_{i=1}^{\infty} \lambda_i |e_i|_V^2 = ||(\lambda_i)||_{\ell^1(\mathbb{N})}.$$

□

Before stating the main result of this section, we need another preliminary result.

Lemma 2.15. *Let $\widetilde{M} \in \mathcal{M}^2(U)$ such that $\langle \widetilde{M}, u\rangle_U$ admits a continuous modification for all $u \in U$. Then $\widetilde{M}$ admits a continuous modification.*

PROOF: Let (f_j) be a complete orthonormal system in U. Define

$$\widetilde{M}_n(t) := \sum_{i=1}^{n} \langle \widetilde{M}, f_j\rangle_U f_j.$$

Then $(\widetilde{M}_n(t))_{n\in\mathbb{N}}$ is a Cauchy sequence in $L^2(\Omega, \mathcal{F}_t; U)$ for $t \geq 0$. Indeed, for positive integers $n > m > N$

$$\mathbb{E}|\widetilde{M}_n(t) - \widetilde{M}_m(t)|_U^2 = \sum_{j=m+1}^{n} \mathbb{E}\langle \widetilde{M}(t), f_j\rangle_U^2 = \sum_{j=m+1}^{n} \langle f_j, \langle\langle \widetilde{M}(t)\rangle\rangle f_j\rangle,$$

which tends to zero as $N \to \infty$ since $\langle\langle M(t)\rangle\rangle \in L_1(U)$. Note furthermore that $\widetilde{M}_n$ admits a continuous modification for all $n \in \mathbb{N}$. Now fix $T > 0$. By Proposition 2.4

$$\mathbb{E} \sup_{t\in[0,T]} |\widetilde{M}_n(t) - \widetilde{M}(t)|) \leq 2 \left(\mathbb{E}|\widetilde{M}_n(T) - \widetilde{M}_m(T)|_U^2\right)^{1/2} \to 0 \quad \text{as } n \to \infty.$$

We have

$$\lim_{k\to\infty} \sup_{t\in[0,T]} |\widetilde{M}_{n_k} - \widetilde{M}| = 0, \quad \text{almost surely,}$$

for a subsequence (n_k). Hence $\widetilde{M}$ is continuous almost surely. □

We can now prove the following result.

Theorem 2.16. *Suppose M is a cylindrical martingale on V of stationary covariance R. Let U be a separable Hilbert space and suppose there exists a linear mapping $J : V \to U$. Suppose that $Q := JRJ^* \in L_1(U)$. Then there exists a unique $\widetilde{M} \in \mathcal{M}^2(U)$ of stationary covariance Q such that*

$$M(t)J^*\varphi = \varphi(\widetilde{M}(t)), \quad t \geq 0, \varphi \in U^*.$$

In particular, $\widetilde{M}$ has a cadlag version. If $M(\cdot)x$ admits a continuous modification for all $x \in V$, then $\widetilde{M}$ admits a continuous modification.

Proof: For every $t \geq 0$ we can define the square integrable random variable $\widetilde{M}(t)$ as in Proposition 2.13. It follows immediately that $\widetilde{M}$ is a square integrable martingale with stationary covariance $Q = JRJ^*$. Since

$$\mathbb{P}(|\widetilde{M}(t) - \widetilde{M}(t_0)| > \varepsilon) \leq \frac{1}{\varepsilon^2}\mathbb{E}|\widetilde{M}(t) - \widetilde{M}(t_0)|^2 = \frac{1}{\varepsilon^2}|t - t_0| \operatorname{tr} Q,$$

we find that $\widetilde{M}$ is stochastically continuous and hence, by Proposition 2.4 admits a cadlag modification.

Suppose now that $M(\cdot)x$ has a continuous version for all $x \in V$. The continuity of $\widetilde{M}$ now follows immediately from Proposition 2.15. □

Remark 2.17. Note that if $R \in L_1(V)$, then M corresponds to a square integrable martingale $\widetilde{M}$ on V. This is easily seen by letting $J : V \to V$ denote the identity mapping. Conversely, to any square integrable martingale $\widetilde{M}$ in U we may associate a cylindrical martingale M on U by defining $M(t)u := \langle \widetilde{M}(t), u\rangle$. Hence cylindrical martingales form a natural generalization of square integrable martingales.

The following result is now an immediate corollary of Lévy's theorem (see [75], Theorem IV.33.1)

Proposition 2.18. *Suppose M is a continuous cylindrical martingale of stationary covariance equal to the identity mapping and with $M(0) = 0$. Then M is a cylindrical Wiener process.*

2.3.1 Cylindrical Lévy process

We call a cylindrical martingale L on V a *cylindrical Lévy process* if Lx is a Lévy process in $\mathbb{R}$ for all $x \in V$ and if L has *independent increments* in the sense that, for all $x \in V$ and $0 \le s \le t$, the increment $(L(t) - L(s))x$ is independent of $\mathcal{F}_s$. Assume for simplicity that $\mathbb{E}L(0)x = 0$ for all $x \in V$ and note that it follows that $\mathbb{E}L(t)x = 0$ for all $x \in V$ and $t \ge 0$. If L is a 2-cylindrical Lévy process then it has stationary covariance and we define the covariance operator R of L as usual by

$$\langle Rx, y\rangle = \mathbb{E}(L(1)x)(L(1)y), \quad x, y \in V.$$

Proposition 2.19. *Suppose L is a cylindrical Lévy process on V with covariance operator R. Let U be a separable Hilbert space and suppose there exists a bounded linear mapping $J : V \to U$, such that $Q := JRJ^* \in L_1(U)$. Let $\widetilde{L} \in \mathcal{M}^2(U)$ be as given by Theorem 2.16 with covariance operator Q.*

Then $\widetilde{L}$ is a Lévy process.

Proof: We only need to verify independence of increments. Since (see [24], Proposition 1.3) $\mathcal{B}(U)$ is generated by sets of the form

$$\{u \in U : \varphi(u) \le \alpha\}, \quad \varphi \in U^*, \alpha \in \mathbb{R},$$

it follows that

$$\mathcal{K}_i := \left\{\varphi_j(\widetilde{L}(t_{i+1}) - \widetilde{L}(t_i)) \le \alpha_j, (\varphi_j)_{j=1}^n \subset U^*, (\alpha_j)_{j=1}^n \subset \mathbb{R}\right\}$$

is a π-system for the σ-algebra generated by the increment $\widetilde{L}(t_{i+1}) - \widetilde{L}(t_i)$. By the assumption of independent increments of L, for $0 \le t_1 \le t_2 \le \ldots \le t_m$, the π-systems $(\mathcal{K}_i)_{i=1}^{m-1}$ are independent, so the generated σ-algebras $\sigma(\mathcal{K}_i)$ are independent or, equivalently, $\widetilde{L}$ has independent increments. Together with the stochastic continuity of $\widetilde{L}$ and the stationarity of the laws $(L(t) - L(s))x$, $0 \le s \le t, x \in V$, this shows that $\widetilde{L}$ is a Lévy process. □

2.3.2 Stochastic integral with respect to cylindrical martingales

We will now define the stochastic integral with respect to cylindrical martingales of stationary covariance. Assume that M is a cylindrical martingale in V of stationary covariance R. Let $J \in L_{\mathrm{HS}}(V;U)$ be an injective mapping (as constructed, for example, by Proposition 2.14). At this point we stress that the injectivity of J will ensure the uniqueness of the definition of the stochastic integral. Let $\widetilde{M}$ be the square integrable martingale defined by Theorem 2.16, of stationary covariance $Q = JRJ^*$. Let $T > 0$. We call a stochastic process Ψ *M-integrable* on $[0,T]$ if

(i) Ψ is a predictable process in the space of linear operators from $R^{1/2}(V)$ into H,

(ii) For almost all $0 \le t \le T$ we have $\Psi(t) \circ R \circ \Psi^*(t) \in L_1(H)$ almost surely, and

(iii)
$$\mathbb{E} \int_0^T \operatorname{tr}[\Psi \circ R \circ \Psi^*] \; dt = \mathbb{E} \int_0^T ||\Psi \circ R^{1/2}||^2_{L_{\mathrm{HS}}(V;H)} \; dt < \infty.$$

The set of M-integrable process on $[0,T]$ is again denoted by $L^2_{M,T}(H)$.

For these processes the *stochastic integral* is defined by

$$\int_0^t \Psi \; dM := \int_0^t \widetilde{\Psi} \; d\widetilde{M}, \quad 0 \le t \le T, \tag{2.9}$$

where

$$\widetilde{\Psi}(t) := \Psi(t) \circ J^{-1}, \quad t \ge 0,$$

is a process in $L^2_{\widetilde{M},T}(H)$ as we will show now.

Theorem 2.20. *The stochastic integral defined by (2.9) is welldefined. In particular, $\widetilde{\Psi} \in L^2_{\widetilde{M},T}(H)$ and the value of $\int_0^t \Psi \; dM$ does not depend on the choice of the space U and the mapping $J \in L_{\mathrm{HS}}(V;U)$. The following version of the Itô isometry holds:*

$$\mathbb{E} \left| \int_0^T \Psi \; dM \right|_H^2 = \mathbb{E} \int_0^T ||\Psi \circ R^{1/2}||^2_{L_{\mathrm{HS}}(V;H)} \; dt, \quad 0 \le t \le T. \tag{2.10}$$

Furthermore all the properties of the stochastic integral of Theorem 2.9, Proposition 2.11 and Proposition 2.12 carry over to this case.

Proof: We define a process Φ with values in $L_{\mathrm{HS}}(U;H)$ as follows. For $u \in U$, let

$$\Phi(t)u := \Psi(t) J^{-1} Q^{1/2} u, \quad t \ge 0,$$

where $J^{-1} : \operatorname{im} J \to V$ is the pseudo-inverse of J.

Note that $Q = (JR^{1/2})(JR^{1/2})^*$. Hence (see [24], Corollary B.4) $\operatorname{im} Q^{1/2} = \operatorname{im} JR^{1/2}$, so that $J^{-1}Q^{1/2} : U \to \operatorname{im} R^{1/2}$. We may conclude that Φ is a welldefined process in the space of linear operators from U into H.

It remains to check that Φ is a process in the space of Hilbert-Schmidt operators $L_{\mathrm{HS}}(U;H)$. To this end, note that, for (f_j) a complete orthonormal system in U,

$$\begin{aligned}\sum_{i=1}^{n}\langle\Phi(t)f_i,\Phi(t)f_i\rangle &= \sum_{i=1}^{n}\langle\Psi(t)J^{-1}Q^{1/2}f_i,\Psi(t)J^{-1}Q^{1/2}f_i\rangle\\ &= \sum_{i=1}^{n}\langle Q^{1/2}(J^{-1})^*\Psi^*(t)\Psi(t)J^{-1}Q^{1/2}f_i,f_i\rangle\\ &= \operatorname{tr}\ Q^{1/2}(J^{-1})^*\Psi^*(t)\Psi(t)J^{-1}Q^{1/2}\\ &= \operatorname{tr}\ \Psi(t)J^{-1}Q(J^{-1})^*\Psi^*(t)\\ &= \operatorname{tr}\ \Psi(t)R\Psi^*(t).\end{aligned}$$

Now for $\widetilde{\Psi}$, with values in the space of linear operators from $Q^{1/2}(U)$ into H, we have

$$\widetilde{\Psi}(t) = \Psi(t)J^{-1} = \Phi(t)Q^{-1/2}, \quad 0\le t\le T.$$

By Theorem 2.7 we find that $\widetilde{\Psi}\in L^2_{\widetilde{M},T}(H)$. We can now define the stochastic integral of Ψ with respect to M by

$$\int_0^t \Psi\ dM := \int_0^T \widetilde{\Psi}\ d\widetilde{M}, \quad 0\le t\le T.$$

Since J is injective, we have $J^{-1}J = I$. Hence for $\varphi\in H^*$ and $\Psi\in L(U;H)$ we have

$$\begin{aligned}\varphi(\widetilde{\Psi}\circ\widetilde{M}(t)) &= \widetilde{\Psi}^*\varphi(\widetilde{M}(t)) = M(t)J^*\widetilde{\Psi}^*\varphi\\ &= M(t)J^*(J^{-1})^*\Psi^*\varphi = M(t)\Psi^*\varphi\\ &= \varphi(\Psi\circ M(t)).\end{aligned}$$

This shows that for simple functions the definition of the stochastic integral is independent of the particular choice of the space U and the mapping $J: V\to U$, which extends to all $\Psi\in L_{M,T}(H)$.

Using (2.7) we have the following version of the Itô isometry:

$$\begin{aligned}\mathbb{E}\left|\int_0^T \Psi\ dM\right|_H^2 = \mathbb{E}\left|\int_0^T \widetilde{\Psi}\ d\widetilde{M}\right|_H^2 &= \mathbb{E}\int_0^T ||\Phi(t)||^2_{L_{\mathrm{HS}}(U;H)}\ dt\\ &= \mathbb{E}\int_0^T ||\Psi\circ R^{1/2}||^2_{L_{\mathrm{HS}}(U;H)}\ dt.\end{aligned}$$

All the mentioned properties follow directly from definition (2.9). □

2.3.3 Reproducing kernel Hilbert space

Let M be a cylindrical martingale in a separable Hilbert space V of injective stationary covariance $R \in L(V)$. Let $\mathcal{H} = \text{im} \quad R^{1/2}$ and equip $\mathcal{H}$ with the inner product

$$\langle R^{1/2}x, R^{1/2}y\rangle_{\mathcal{H}} := \langle x, y\rangle_V.$$

Then $\mathcal{H}$ is a Hilbert space, called the *reproducing kernel Hilbert space* (*RKHS*) corresponding to M. We can rephrase our main result on stochastic integration in terms of $\mathcal{H}$.

Theorem 2.21. *Let M be a cylindrical martingale of stationary covariance with RKHS $\mathcal{H}$. Then*

$$L^2_{M,T}(H) = L^2(\Omega \times [0,T], \mathcal{P}_T, d\mathbb{P} \otimes dt; L_{\text{HS}}(\mathcal{H}; H)). \tag{2.11}$$

and for $\Psi \in L^2_{M,T}(H)$ we have

$$\mathbb{E}\left|\int_0^T \Psi \, dM\right|_H^2 = \mathbb{E}\int_0^T ||\Psi||^2_{L_{\text{HS}}(\mathcal{H};H)} \, dt. \tag{2.12}$$

2.4 Examples

In this section we provide some interesting examples of noise processes. For this it will be convenient to have the concept of a random measure at our disposal.

Definition 2.22. *Let $(S, \mathcal{S})$ be a measurable space. A* random measure *ξ on S is a mapping $\xi : \Omega \times \mathcal{S} \to [0, \infty]$ such that $\xi(\omega, A)$ is measurable in $\omega \in \Omega$ for all $A \in \mathcal{S}$, and a σ-finite measure in $A \in \mathcal{S}$ for all $\omega \in \Omega$.*

Note that $\xi(\cdot, A)$ is a $[0, \infty]$-valued random variable for every $A \in \mathcal{S}$.

2.4.1 Space-time Brownian motion

Let $b > 0$, $d \in \mathbb{N}$ and let $\mathcal{O} = [0, b]^d \subset \mathbb{R}^d$.

A measurable mapping $B : \Omega \times \mathbb{R}_+ \times \mathcal{O} \to \mathbb{R}$ is a *space-time Brownian motion* if $\{B(t, x) : (t, x) \in \mathbb{R}_+ \times \mathcal{O}\}$ is a collection of Gaussian random variables such that

(i) $\mathbb{E}B(t, x) = 0$ for all $(t, x) \in \mathbb{R}_+ \times \mathcal{O}$, and

(ii) $\mathbb{E}B(t, x)B(s, y) = \text{Leb}(A_{t,x} \cap A_{s,y})$ for all $(t, x), (s, y) \in \mathbb{R}_+ \times \mathcal{O}$.

Here for $(t,x) \in \mathbb{R}_+ \times \mathcal{O}$ the sets $A_{t,x}$ are defined by

$$A_{t,x} := \{(s,y) \in \mathbb{R}_+ \times \mathcal{O} : 0 \leq s \leq t, 0 \leq y_i \leq x_i, i = 1, \ldots, d\},$$

and Leb denotes Lebesgue measure on $\mathbb{R}^d$.

We define a random measure $\mathcal{W}$ on $\mathbb{R}_+ \times \mathcal{O}$ by

$$\mathcal{W}(\cdot, A_{t,x}) := B(t,x), \quad (t,x) \in \mathbb{R}_+ \times \mathcal{O}.$$

Then $\mathcal{W}$ is a random measure called a *white noise random measure* on $\mathbb{R}_+ \times \mathcal{O}$.

Let $f \in L^2(\mathcal{O}) = L^2(\mathcal{O}, \mathrm{Leb})$ and define a mapping

$$W : [0,\infty) \to L(L^2(\mathcal{O}); L^2(\Omega))$$

by

$$W(t)f := \int_{\mathcal{O}} \int_0^t f(x)\, d\mathcal{W}(s,x), \quad t \geq 0. \tag{2.13}$$

We have the following relation between a cylindrical Wiener process and a white noise random measure ([43], Theorem 3.2.4).

Theorem 2.23. *Let $\mathcal{H} = L^2(\mathcal{O})$.*

(i) Let $\mathcal{W}$ be a white noise random measure on $\mathbb{R}_+ \times \mathcal{O}$. Then W defined by (2.13) is a cylindrical Wiener process with RKHS $\mathcal{H}$.

(ii) Suppose $\widetilde{W}$ is a cylindrical Wiener process with RKHS $\mathcal{H}$. Then there exists a white noise random measure $\mathcal{W}$ on $\mathbb{R}_+ \times \mathcal{O}$ such that W constructed in (2.13) has the property

$$W(t)f = \widetilde{W}(t)f, \quad \text{almost surely for all } f \in L^2(\mathcal{O}).$$

2.4.2 Environmental pollution

Consider an environment which suffers from pollution at random times, in random amounts.

As environment choose $\mathcal{O} \subset \mathbb{R}^d$. At increasing random times $(\tau_k)_{k=1}^{\infty}$ pollution enters the environment, of random magnitudes (σ_k) and at random locations $(\xi_k) \subset \mathcal{O}$. All ξ_k, σ_k are independent identically distributed random variables with probability laws μ, ν on $\mathcal{O}$ and $[0,\infty)$, respectively, both with finite second moments, and independent of the random times τ_k, for which $\tau_k - \tau_{k-1}$ is exponentially distributed with rate λ, $k = 1, \ldots, \infty$, $\tau_0 = 0$.

Let $(Z(t))$ denote the measure valued compound Poisson process describing cumulative pollution. Then

$$Z(t) = \sum_{\tau_k \leq t} \sigma_k \delta_{\xi_k}, \quad t \geq 0,$$

with δ_ξ denoting the Dirac delta function at $\xi \in \mathcal{O}$. Note that

$$\mathbb{E}Z(t)(A) = \lambda t \int_0^\infty \sigma \, d\nu(\sigma) \int_{\mathcal{O}} \delta_\xi(A) \, d\mu(\xi) = \lambda t \int_0^\infty \sigma \, d\nu(\sigma)\mu(A),$$

for $A \in \mathcal{B}(\mathcal{O})$.

Let $\widetilde{Z}(t) = Z(t) - \mathbb{E}Z(t)$. We may interpret $\widetilde{Z}$ as a cylindrical Lévy process on $L^2(\mathcal{O}, \mu)$ as follows. For $\psi \in L^2(\mathcal{O}, \mu)$ let

$$M(t)\psi := \int_{\mathcal{O}} \psi \, d\widetilde{Z}(t).$$

It can be calculated that

$$\begin{aligned} &\mathbb{E}|M(1)\psi|^2 \\ &= (\lambda^2 + \lambda) \int_0^\infty \sigma^2 \, d\nu(\sigma) \int_{\mathcal{O}} \psi^2 \, d\mu - \lambda^2 \left(\int_0^\infty \sigma \, d\nu(\sigma) \right)^2 \left(\int_{\mathcal{O}} \psi \, d\mu \right)^2. \end{aligned}$$

So as long as $\int_0^\infty \sigma^2 \, d\nu(\sigma) < \infty$ we have that M is a cylindrical martingale. Furthermore, it follows that the covariance operator R is given by

$$R\psi = (\lambda^2 + \lambda) \int_0^\infty \sigma^2 \, d\nu(\sigma)\psi - \lambda^2 \left(\int_0^\infty \sigma \, d\nu(\sigma) \right)^2 \langle \psi, \mathbb{1} \rangle_{L^2(\mathcal{O},\mu)} \mathbb{1},$$

for $\psi \in L^2(\mathcal{O}, \mu)$, with $\mathbb{1}$ denoting the constant function equal to 1, μ-almost everywhere on $\mathcal{O}$. By virtue of the fact that ν and μ are probability measures, we may estimate

$$\langle \psi, R\psi \rangle \geq \lambda \int_0^\infty \sigma^2 d\nu(\sigma) \int_{\mathcal{O}} \psi^2 \, d\mu.$$

It follows that R is a coercive mapping on $L^2(\mathcal{O}, \mu)$. Therefore the RKHS $\mathcal{H}$ of M is isomorphic to $L^2(\mathcal{O}; \mu)$; however the natural inclusion is not an isometry.

2.5 Stochastic differential equations

To study stochastic differential equations in H, we use the following ingredients.

Hypothesis 2.24. *(i) M is a cylindrical martingale of stationary covariance with RKHS $\mathcal{H}$.*

(ii) $(S(t))_{t \geq 0}$ is a strongly continuous semigroup in H;

(iii) $F : \mathfrak{D}(F) \to H$ is a mapping such that $\mathfrak{D}(F)$ is dense in H and there exists a function $a \in L^1_{\mathrm{loc}}([0, \infty))$ such that for all $t > 0$ and $x, y \in \mathfrak{D}(F)$,

$$\begin{aligned} |S(t)F(x) - S(t)F(y)|_H &\leq a(t)|x - y|_H \quad \textit{and} \\ |S(t)F(x)|_H &\leq a(t) \, (1 + |x|_H); \end{aligned}$$

(iv) $G : \mathfrak{D}(G) \to L_{\mathrm{HS}}(\mathcal{H}; H)$ *is a mapping such that* $\mathfrak{D}(G)$ *is dense in* H *and there exists a function* $b \in L^2_{\mathrm{loc}}([0,\infty))$ *such that for all* $t > 0$ *and* $x, y \in \mathfrak{D}(G)$,

$$||S(t)G(x) - S(t)G(y)||_{L_{\mathrm{HS}}(\mathcal{H};H)} \leq b(t)|x-y|_H \quad \textit{and} \\ ||S(t)G(x)||_{L_{\mathrm{HS}}(\mathcal{H};H)} \leq b(t)(1+|x|_H).$$

Remark 2.25. Note that (iii) and (iv) of Hypothesis 2.24 are trivially implied by Lipschitz continuity of F and G.

A *stochastic differential equation* or *stochastic Cauchy problem* is an equation of the form

$$\begin{cases} dX = [AX + F(X)]\ dt + G(X)\ dM, & t \geq t_0 \\ X(t_0) = X_0. \end{cases} \tag{2.14}$$

Equation (2.14) is only a formal expression. Below we define the meaning of a solution to (2.14):

Definition 2.26. *Suppose Hypothesis 2.24 holds. Let* $X_0 \in L^2(\Omega, \mathcal{F}_{t_0}; H)$*. A* (mild) solution to (2.14) starting at time t_0 from X_0 *is a predictable process* $X = (X(t))_{t \geq t_0}$ *such that*

$$\sup_{t\in[t_0,T]} \mathbb{E}|X(t)|_H^2 < \infty, \quad \textit{for all } T > t_0,$$

satisfying

$$\begin{aligned} X(t) = S(t-t_0)X_0 &+ \int_{t_0}^t S(t-s)F(X(s))\ ds \\ &+ \int_{t_0}^t S(t-s)G(X(s))\ dM(s), \quad t \geq t_0. \end{aligned} \tag{2.15}$$

If in addition $X(t) \in \mathfrak{D}(A)$ *almost surely for* $t \geq t_0$*, and*

$$X(t) = X_0 + \int_{t_0}^t AX(s) + F(X(s))\ ds + \int_{t_0}^t G(X(s))\ dM(s), \quad t \geq t_0, \tag{2.16}$$

then X *is called a* strong solution to (2.14).

The following result on existence and uniqueness of solutions is classical (see [71], Theorem 9.29).

Theorem 2.27. *Suppose Hypothesis 2.24 holds. Then for all* $t_0 \geq 0$ *and* $X_0 \in L^2(\Omega, \mathcal{F}_{t_0}; H)$ *there exists a unique (predictable) solution* $X = (X(t; t_0, X_0))_{t \geq t_0}$ *to (2.14). Furthermore for all* $t_0 < T < \infty$ *there exists a constant* $L > 0$ *such that for all* $x, y \in H$

$$\sup_{t\in[t_0,T]} \mathbb{E}|X(t;t_0,x) - X(t;t_0,y)|_H^2 \leq L|x-y|_H^2 \\ \sup_{t\in[t_0,T]} \mathbb{E}|X(t;t_0,x)|_H^2 \leq L(1+|x|_H^2).$$

Finally for all $0 \leq t_0 \leq t$ and $x \in H$ the law of $X(t; t_0, x)$ does not depend on the choice of the probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$.

The following proposition ([39], Proposition 1.3 (i), [24], Proposition 7.3) is important in the study of regularity of solutions.

Definition 2.28. *A strongly continuous semigroup $(S(t))_{t \geq 0}$ on a Banach space X is called a* generalized contraction *if $||S(t)x|| \leq \mathrm{e}^{\omega t}||x||$ for some $\omega \in \mathbb{R}$ and all $t \geq 0$, $x \in X$.*

Proposition 2.29 (Estimate for stochastic convolution)**.** *Consider the process*

$$M_A^\Phi(t) := \int_0^t S(t-s)\Phi(s)\ dM(s),$$

where M is a cylindrical martingale of stationary covariance with RKHS $\mathcal{H}$, $(S(t))_{t \geq 0}$ is a strongly continuous semigroup on H, and $\Phi \in L^2_{M,T}(H)$.

(i) Suppose that S is a generalized contraction semigroup and $p \in (0, 2]$. Then there exists a constant $K_{p,T}$ such that

$$\mathbb{E}\left[\sup_{t \in [0,T]} |M_A^\Phi(t)|^p\right] \leq K_{p,T}\mathbb{E}\left(\int_0^T ||\Phi(t)||^2_{L_{\mathrm{HS}}(\mathcal{H};H)}\ dt\right)^{p/2}.$$

(ii) Suppose that $p > 2$, M is continuous. Then there exists a constant $K_{p,T}$ such that

$$\mathbb{E}\left[\sup_{t \in [0,T]} |M_A^\Phi(t)|^p\right] \leq K_{p,T}\mathbb{E}\int_0^T ||\Phi(t)||^p_{L_{\mathrm{HS}}(\mathcal{H};H)}\ dt$$

for all $\Phi \in L^2_{M,T}(H)$ for which $E\int_0^T ||\Phi||^p_{L_{\mathrm{HS}}(\mathcal{H};H)}\ dt < \infty$. Furthermore

$$\lim_{n\to\infty} \mathbb{E} \sup_{t \in [0,T]} |M_A^\Phi(t) - M_{A_n}^\Phi(t)|^p = 0,$$

where $(A_n)_{n \in \mathbb{N}}$ are approximations of A (see section 6.4.2).

The constant $K_{p,T}$ depends only on p, T,

Remark 2.30. For the case where S is a generalized contraction, note that we may assume without loss of generality that S is a contraction. This case is proven in [39], Proposition 1.3 (i). Their proof is based on Szekőfalvi-Nagy's theorem on unitary dilations ([27], Theorem 7.2.1).

The second case is proven, using a factorization of the stochastic integral, in [24], Proposition 7.3 for M a cylindrical Wiener process and $(A_n)_{n \in \mathbb{N}}$ Yosida approximants. The proof extends without problem to our case.

For one case we can establish the cadlag property for paths of solutions (see [71], Theorem 9.29).

Hypothesis 2.31. *(i) M is a cylindrical martingale of stationary covariance with RKHS $\mathcal{H}$;*

(ii) $(S(t))_{t\geq 0}$ is a generalized contraction;

(iii) F satisfies the conditions of Hypothesis 2.24;

(iv) G is globally Lipschitz.

Theorem 2.32. *Suppose Hypothesis 2.31 holds. Then the solution $X(\cdot; t_0, X_0)$ of* (2.14) *has a cadlag version.*

Another important case is when we can exploit the *factorization method* ([23], [71], Chapter 11). For this we introduce a third set of assumptions on M, S, F and G.

Hypothesis 2.33. *(i) M is a continuous cylindrical martingale with RKHS $\mathcal{H}$;*

(ii) $(S(t))_{t\geq 0}$ is a strongly continuous semigroup acting on H;

(iii) Fix $q > 2$ and $\alpha \in (1/q, 1/2)$. There exists a measurable function $a : (0, T] \to (0, \infty)$ such that

$$\int_0^T a(t)t^{-\alpha}\, dt < \infty,$$

and for all $t \in (0, T]$ and $x, y \in H$,

$$\begin{aligned} |S(t)F(x)|_H &\leq a(t)(1 + |x|_H) \\ |S(t)(F(x) - F(y))|_H &\leq a(t)|x - y|_H. \end{aligned}$$

(iv) There exists a measurable function $b : (0, T] \to (0, \infty)$ such that

$$\int_0^T b^2(t)t^{-2\alpha}\, dt < \infty,$$

and for all $t \in (0, T]$ and $x, y \in H$,

$$\begin{aligned} ||S(t)G(x)||_{L_{\mathrm{HS}}(\mathcal{H};H)} &\leq b(t)(1 + |x|_H), \\ ||S(t)(G(x) - G(y))||_{L_{\mathrm{HS}}(\mathcal{H};H)} &\leq b(t)|x - y|_H. \end{aligned}$$

(v) $X_0 \in L^q(\Omega; H)$.

Note that (i), (iii) and (iv) of Hypothesis 2.33 are stronger than (i), (iii) and (iv) of Hypothesis 2.24. For this situation we can establish continuity of paths (see [71], Theorem 11.8).

Theorem 2.34. *Suppose Hypothesis 2.33 holds. Then there exists a modification of the solution* $X(\cdot; t_0, X_0)$ *of (2.14) such that* $X \in L^q(\Omega; C([0,T]; H))$*. In particular,* X *has continuous trajectories.*

Remark 2.35. The proof in [71], Theorem 11.8, for the Wiener case can be extended without problems to the case of continuous cylindrical martingales, using the BDG inequalities (2.2).

2.6 Notes and remarks

The results of this chapter are not new. For finite dimensional stochastic integration and differential equations, see [44] and [73]. The Hilbert space theory is developed in [24] (for Wiener processes) and [43], [61] and [71] (for more general processes).

In this thesis we concentrate on stochastic differential equations with Hilbert space valued solutions. It is essentially the Itô isometry which makes Hilbert spaces tailormade for stochastic integration theory.

It is however possible to work with Banach space valued stochastic integration and stochastic differential equations. This theory has been developed to a great extent in recent years, but is rather sophisticated on the functional analytical level. Furthermore, for our prime example, the stochastic delay equation, the natural Banach space to work with is the space of continuous functions. This Banach space does not possess the so-called *UMD-property*, which makes it an inconvenient space for stochastic analysis. The interested reader is referred to [91], [93] or [88].

The example of environmental pollution of Section 2.4.2 is based on similar examples in [71] and [43]. For another discussion of the dynamics of pollution, see [86].

CHAPTER

THREE

Stochastic differential equations with delay

Differential equations with delayed dependence have been studied for several centuries now. See for example an overview article by Erhard Schmidt [78]. One can think of both ordinary differential equations and partial differential equations, perturbed by a dependence on the 'past' of the process, leading to *functional differential equations* (see [28], [29], [38]) and *partial functional differential equations* (see [82], [83]), respectively. The latter class has attracted some amount of research activity recently, see for example [53] and [94].

Strictly speaking a delay differential equation is a specific example of a functional differential equation, in which the functional part of the differential equation is the evaluation of a functional on the past of the process. However, we will freely interchange the terminology 'delay differential equation' and 'functional differential equation'.

We wish to extend the analysis of deterministic delay differential equations to delay differential equations perturbed by random noise or stochastic delay differential equations.

In Section 3.1 we show how to study deterministic delay equations in a Hilbert space context, introducing the delay semigroup. In Section 3.2 we take advantage of this by explaining how to study stochastic delay differential equations. Examples of such evolutions are described in Sections 3.3 (stochastic functional differential equations) and 3.4 (stochastic partial differential equations with delay).

This chapter concludes with two useful results on the delay semigroup. In Section 3.5 we show that for a large amount of cases the delay semigroup is eventually compact. Section 3.6 explains how, in quite some cases, we may renorm the state space in such a way that the delay semigroup becomes a contraction semigroup.

3.1 Abstract delay differential equations

We present here the abstract framework for the study of deterministic delay differential equations, or *delay equation*, of [7]. Here we choose to work mainly in L^2 spaces since these spaces are particularly suited for stochastic analysis (see the discussion in Section 2.6). Deterministic delay equations may also be studied in spaces of continuous functions, see [29] and [32], Section VI.6.

Let X, Z be Banach spaces, and let $W^{1,p}([-1,0];Z)$ denote the *Sobolev space* consisting of equivalence classes of functions in $L^p([-1,0];Z)$ which have a weak derivative in $L^p([-1,0;V)$ (see [81], Chapter 4).

Consider the abstract differential equation with delay

$$\begin{cases} \frac{d}{dt}u(t) = Bu(t) + \Phi u_t, \quad t > 0, \\ u(0) = x, \\ u_0 = f, \end{cases} \tag{3.1}$$

under the following assumptions:

Hypothesis 3.1. *(i) $x \in X$;*

(ii) B is the generator of a strongly continuous semigroup $(S(t))_{t\geq 0}$ in X;

(iii) $\mathfrak{D}(B) \hookrightarrow Z \hookrightarrow X$ with the injections being dense;

(iv) $f \in L^p([-1,0];Z)$, $1 \leq p < \infty$;

(v) $\Phi : W^{1,p}([-1,0];Z) \to X$ is a bounded linear operator.

(vi) $u : [-1,\infty) \to X$ and for $t \geq 0$, $u_t : [-1,0] \to X$ is defined by $u_t(\sigma) = u(t+\sigma)$, $\sigma \in [-1,0]$.

In general, if $(\xi(t))_{t\in[-1,\infty)}$ is a process, then the process $(\xi_t)_{t\geq 0}$ with values in a function space, defined by $\xi_t(\sigma) := \xi(t+\sigma)$, $t \geq 0$, $\sigma \in [-1,0]$, is called the *segment process* of ξ. So here $(u_t)_{t\geq 0}$ is the segment process of $(u(t))_{t\in[-1,\infty)}$. It keeps track of the history of $(u(t))_{t\in[-1,\infty)}$.

Definition 3.2. *A* classical solution *of (3.1) is a function $u : [-1,\infty) \to X$ that satisfies*

(i) $u \in C([-1,\infty);X) \cap C^1([0,\infty);X)$;

(ii) $u(t) \in \mathfrak{D}(B)$ and $u_t \in W^{1,p}([-1,0];Z)$ for all $t \geq 0$;

(iii) u satisfies (3.1).

To employ a semigroup approach we introduce the Banach space

$$\mathcal{E}^p := X \times L^p([-1,0];Z),$$

and the closed, densely defined operator in $\mathcal{E}^p$,

$$A := \begin{bmatrix} B & \Phi \\ 0 & \frac{d}{d\sigma} \end{bmatrix}, \quad \mathfrak{D}(A) = \left\{ \begin{pmatrix} x \\ f \end{pmatrix} \in \mathfrak{D}(B) \times W^{1,p}([-1,0];Z) : f(0) = x \right\}. \tag{3.2}$$

The equation (3.1) is called *well-posed* if for all $(x,f) \in \mathfrak{D}(A)$, there exists a unique classical solution of (3.1) that depends continuously on the initial data (in the sense of uniform convergence on compact intervals).

It is shown in [7], Corollary 3.7, that A generates a strongly continuous semigroup in $\mathcal{E}^p$ if and only if (3.1) is well-posed. In this case the semigroup is called an *(abstract) delay semigroup*.

Furthermore, sufficient conditions on Φ are given for this to be the case:

Hypothesis 3.3. *Let $S_t : X \to L^p([-1,0];Z)$ be defined by*

$$(S_t x)(\tau) := \begin{cases} S(t+\tau)x & \text{if } -t < \tau \leq 0, \\ 0 & \text{if } -1 \leq \tau \leq -t, \end{cases} \qquad t \geq 0.$$

Let $(T_0(t))_{t\geq 0}$ be the nilpotent left shift semigroup on $L^p([-1,0];Z)$. Assume that there exists a function $q : [0,\infty) \to [0,\infty)$ with $\lim_{t\downarrow 0} q(t) = 0$, such that

$$\int_0^t ||\Phi(S_r x + T_0(r)f)||\, dr \leq q(t) \left|\left| \begin{pmatrix} x \\ f \end{pmatrix} \right|\right| \tag{3.3}$$

for all $t > 0$ and $\begin{pmatrix} x \\ f \end{pmatrix} \in \mathfrak{D}(A)$. Furthermore suppose that either

(A) $Z = X$ *or*

(B) *(i)* $(B, \mathfrak{D}(B))$ *generates an analytic semigroup $(S(t))_{t\geq 0}$ on X, and*

(ii) for some $\delta > \omega_0(B)$ there exists $\vartheta < \frac{1}{p}$ such that

$$\mathfrak{D}((-B+\delta I)^{\vartheta}) \hookrightarrow Z \hookrightarrow X,$$

with the injections being dense.

Theorem 3.4. *Assume Hypothesis 3.3 holds. Then $(A, \mathfrak{D}(A))$ is the generator of a strongly continuous semigroup $(T(t))_{t\geq 0}$ on $\mathcal{E}^p$.*

PROOF: See [7], Theorem 3.26 and Theorem 3.34. □

Remark 3.5. Condition (3.3) is slightly stronger than needed but will also provide us with sufficient regularity of the delay semigroup later on (see [7], Section 4.1).

Example 3.6. Let $\Phi : C([-1,0];Z) \to X$ be given by

$$\Phi(f) := \int_{-1}^{0} d\eta\ f,$$

where $\eta : [-1,0] \to L(Z;X)$ is of bounded variation.

Suppose that either $Z = X$ or $(B, \mathfrak{D}(B))$ satisfies the assumptions (B-i) and (B-ii) of Hypothesis 3.3. Then the conditions of Hypothesis 3.3 are satisfied and hence $(A, \mathfrak{D}(A))$ generates a strongly continuous semigroup (see [7], Theorem 3.29 and Theorem 3.35). $\diamond$

3.2 Stochastic differential equations with delay

In the previous section we explained how to interpret a delay differential equation from an infinite-dimensional point of view. How to do this if we perturb such an equation by noise (which possibly also depends on the history of the process)?

Consider the following general form of a *stochastic delay differential equation*.

$$\begin{cases} dY(t) = [BY(t) + \Phi Y_t + \varphi(Y(t), Y_t)]\ dt + \psi(Y(t), Y_t)\ dM(t), \\ \hfill t \geq 0, \\ Y(0) = x, \\ Y(t) = f(t), \hfill -1 \leq t \leq 0. \end{cases} \tag{3.4}$$

where

Hypothesis 3.7. *(i) X and Z are Hilbert spaces;*

(ii) B, Φ, x and f are as in Hypothesis 3.1 with $p = 2$, and Hypothesis 3.3 is satisfied;

(iii) $\varphi : \mathcal{E}^2 \to X$ is Lipschitz;

(iv) M is a cylindrical martingale with RKHS $\mathcal{H}$;

(v) Either

(a) $\psi : \mathcal{E}^2 \to L_{\mathrm{HS}}(\mathcal{H};X)$ is Lipschitz, or

(b) $X = Z$, $\Phi \in L(L^2([-1,0];X);X)$ and $\psi : \mathfrak{D}(\psi) \subset \mathcal{E}^2 \to L(\mathcal{H};X)$ with $\mathfrak{D}(\psi)$ dense in $\mathcal{E}^2$, and, for all $t > 0$,

$$||S(t)\psi(x) - S(t)\psi(y)||_{L_{\mathrm{HS}}(\mathcal{H};X)} \leq q(t)|x-y|_{\mathcal{E}^2}, \quad x, y \in \mathfrak{D}(\psi), \quad \textit{and}$$
$$||S(t)\psi(x)||_{L_{\mathrm{HS}}(\mathcal{H};X)} \leq q(t)(1 + |x|_{\mathcal{E}^2}), \quad x \in \mathfrak{D}(\psi),$$

where $(S(t))_{t\geq 0}$ is the strongly continuous semigroup in $L(X)$ generated by B, and with $q \in L^2_{\mathrm{loc}}([0,\infty))$.

Let A, defined by (3.2), generate a delay semigroup $(T(t))_{t\geq 0}$ as in the previous section. Note that, for a large class of noise processes and choices of ψ, a solution $(Y(t))$ to (3.4) has paths which are non-differentiable almost surely, even when the initial condition is smooth. This implies that we can not just rephrase (3.4) into a strong stochastic evolution equation in $\mathcal{E}^2$ of the form

$$d\xi = [A\xi(t) + F(\xi(t))]\, dt + G(\xi(t))\, dM(t),$$

for the process $\xi(t) = \begin{pmatrix} Y(t) \\ Y_t \end{pmatrix}$, since in general $Y_t \notin W^{1,2}([-1,0];Z)$ and hence $\xi(t) \notin \mathfrak{D}(A)$, $t \geq 0$.

Instead we use the mild form

$$\xi(t) = T(t)\begin{pmatrix} x \\ f \end{pmatrix} + \int_0^t T(t-s)F(\xi(s))\, ds + \int_0^t T(t-s)G(\xi(s))\, dM(s), \quad t \geq 0, \tag{3.5}$$

where we define the mappings $F : \mathcal{E}^2 \to \mathcal{E}^2$ and $G : \mathcal{E}^2 \to L_{\mathrm{HS}}(\mathcal{H};\mathcal{E}^2)$ by

$$F\begin{pmatrix} x \\ f \end{pmatrix} := \begin{pmatrix} \varphi(x,f) \\ 0 \end{pmatrix} \quad \text{and} \quad G\begin{pmatrix} x \\ f \end{pmatrix} := \begin{pmatrix} \psi(x,f) \\ 0 \end{pmatrix}, \quad \begin{pmatrix} x \\ f \end{pmatrix} \in \mathcal{E}^2.$$

By its definition F is Lipschitz, and if case (v-a) holds G is Lipschitz as well. We will now show that, also in case (v-b), G satisfies sufficient conditions in order for a solution to (3.5) to exists.

Throughout this chapter, let

$$\begin{aligned} \pi_1 :& X \times L^2([-1,0];Z) \to X \quad \text{and} \\ \pi_2 :& X \times L^2([-1,0];Z) \to L^2([-1,0];Z) \end{aligned}$$

denote the canonical orthogonal projections.

Lemma 3.8. *Suppose condition Hypothesis 3.7, (v-b) holds. Then the mapping $T(t)G$ with G as above satisfies the estimates of Hypothesis 2.24, (iv).*

Proof: Let $(T_0(t))_{t\geq 0}$ be the strongly continuous semigroup in $\mathcal{E}^2$ generated by

$$A_0 = \begin{pmatrix} B & 0 \\ 0 & \frac{d}{d\sigma} \end{pmatrix}, \quad \mathfrak{D}(A_0) = \mathfrak{D}(A).$$

Note that

$$\begin{aligned} \pi_1 T_0(t)G(x) &= S(t)\psi(x), \quad \text{and} \\ \pi_2 T_0(t)G(x)(\sigma) &= \begin{cases} S(t+\sigma)\psi(x), & t+\sigma > 0 \\ 0, & t+\sigma < 0, \end{cases} \end{aligned}$$

for $t > 0$ and $x \in \mathfrak{D}(\psi)$.

It can be verified that, for general $\Psi \in L_{\mathrm{HS}}(\mathcal{H}; \mathcal{E}^2)$,

$$||\Psi||^2_{L_{\mathrm{HS}}(\mathcal{H};\mathcal{E}^2)} \leq ||\pi_1 \Psi||^2_{L_{\mathrm{HS}}(\mathcal{H};X)} + \int_{-1}^0 ||\pi_2 \Psi(\sigma)||^2_{L_{\mathrm{HS}}(\mathcal{H};X)} \, d\sigma.$$

Hence, for $0 \leq t \leq 1$,

$$\begin{aligned}
&||T_0(t)(G(x) - G(y))||^2_{L_{\mathrm{HS}}(\mathcal{H};\mathcal{E}^2)} \\
&= ||S(t)(\psi(x) - \psi(y))||^2_{L_{\mathrm{HS}}(\mathcal{H};X)} + \int_{-t}^0 ||S(t+\sigma)(\psi(x) - \psi(y))||^2_{L_{\mathrm{HS}}(\mathcal{H};X)} \, d\sigma \\
&\leq \left(q(t)^2 + \int_0^t q(r)^2 \, dr \right) |x - y|^2_{\mathcal{E}^2},
\end{aligned}$$

and for $t > 1$

$$\begin{aligned}
&||T_0(t)(G(x) - G(y))||^2_{L_{\mathrm{HS}}(\mathcal{H};\mathcal{E}^2)} \\
&= ||S(t)(\psi(x) - \psi(y))||^2_{L_{\mathrm{HS}}(\mathcal{H};X)} + \int_{-1}^0 ||S(t+\sigma)(\psi(x) - \psi(y))||^2_{L_{\mathrm{HS}}(\mathcal{H};X)} \, d\sigma \\
&\leq \left(q(t)^2 + \int_0^1 q(t+r-1)^2 \, dr \right) |x - y|^2_{\mathcal{E}^2}.
\end{aligned}$$

Note that the mapping

$$t \mapsto m(t) := \begin{cases} q(t)^2 + \int_0^t q(r)^2 \, dr, & 0 \leq t \leq 1, \\ q(t)^2 + \int_0^1 q(t+r-1)^2 \, dr, & t > 1. \end{cases}$$

is locally integrable, by local square integrability of q.

By variation of constants,

$$T(t)G(x) = T_0(t)G(x) + \int_0^t T_0(t-s) \begin{pmatrix} 0 & \Phi \\ 0 & 0 \end{pmatrix} T(s)G(x) \, ds,$$

and therefore, for $0 \leq t \leq T$,

$$\begin{aligned}
&||T(t)(G(x) - G(y))||^2_{L_{\mathrm{HS}}(\mathcal{H};\mathcal{E}^2)} \\
&\leq 2||T_0(t)(G(x) - G(y))||^2_{L_{\mathrm{HS}}(\mathcal{H};\mathcal{E}^2)} + \\
&\quad 2 \left|\left| \int_0^t T_0(t-s) \begin{pmatrix} 0 & \Phi \\ 0 & 0 \end{pmatrix} T(s)(G(x) - G(y)) \, ds \right|\right|^2_{L_{\mathrm{HS}}(\mathcal{H};\mathcal{E}^2)} \\
&\leq 2m(t)|x - y|^2_{\mathcal{E}^2} + 2T \int_0^t \left|\left| T_0(t-s) \begin{pmatrix} 0 & \Phi \\ 0 & 0 \end{pmatrix} T(s)(G(x) - G(y)) \right|\right|^2_{L_{\mathrm{HS}}(\mathcal{H};\mathcal{E}^2)} \, ds \\
&\leq 2m(t)|x - y|^2_{\mathcal{E}^2} + 2TM^2(e^{2\omega T} \vee 1)||\Phi||^2 \int_0^t ||T(s)(G(x) - G(y))||^2_{L_{\mathrm{HS}}(\mathcal{H};\mathcal{E}^2)} \, ds.
\end{aligned}$$

where $||T_0(t)|| \leq M\mathrm{e}^{\omega t}$, $t \geq 0$. Hence, by Gronwall,

$$||T(t)(G(x) - G(y))||^2_{L_{\mathrm{HS}}(\mathcal{H};\mathcal{E}^2)} \leq 2m(t)|x-y|^2_{\mathcal{E}^2} \exp(2TM^2(\mathrm{e}^{2\omega T} \vee 1)||\Phi||^2 t).$$

We see that this mapping is locally integrable in t and hence the first requirement of Hypothesis 2.24 (iv) is satisfied.

In a similar way we may estimate $||T(t)G(x)||_{L_{\mathrm{HS}}(\mathcal{H};\mathcal{E}^2)}$ to obtain the second requirement. □

By Theorem 2.27 there exists a unique solution to (3.5). We will now show that such a solution solves (3.4).

Theorem 3.9. *Suppose $(\xi(t))_{t\geq 0}$ is the unique solution of (3.5). Then $Y(t) := \pi_1\xi(t)$, $t \geq 0$, satisfies (3.4), and $Y_t := \pi_2\xi(t)$, $t \geq 0$ is the segment process of $(Y(t))_{t\geq 0}$.*

PROOF: We prove the first statement. Rewrite (3.5) into

$$\begin{aligned}
\pi_1\xi(t) &= y\left(t; \begin{pmatrix} x \\ f \end{pmatrix}\right) + \int_0^t y(t-s; F(\xi(s))\; ds + \int_0^t y(t-s; G(\xi(s))\; dM(s) \\
\pi_2\xi(t) &= y_t \begin{pmatrix} x \\ f \end{pmatrix} + \int_0^t y_{t-s}(F(\xi(s)))\; ds + \int_0^t y_{t-s}(G(\xi(s)))\; dM(s),
\end{aligned}$$

where, for $\zeta \in \mathcal{E}^2$, $(y(t;\zeta))_{t\geq -1}$ is the solution of

$$\frac{d}{dt}y(t) = By(t) + \Phi y_t, \quad t \geq 0, \quad \text{and} \quad y(0) = \pi_1\zeta, \quad y_0 = \pi_2\zeta,$$

so that $y(t;\zeta) = \pi_1 T(t)\zeta$, and $(y_t(\zeta))_{t\geq 0}$ is the segment process of $(y(t;\zeta))_{t\geq -1}$.

First note that, for almost all $\sigma \in [-1, 0]$ and $t \geq 0$,

$$\begin{aligned}
\pi_2\xi(t)(\sigma) &= y\left(t+\sigma; \begin{pmatrix} x \\ f \end{pmatrix}\right) + \int_0^t y(t+\sigma-s; F(\xi(s)))\; ds \\
&\quad + \int_0^t y(t+\sigma-s; G(\xi(s)))\; dM(s) \\
&= y\left(t+\sigma; \begin{pmatrix} x \\ f \end{pmatrix}\right) + \int_0^{t+\sigma} y(t+\sigma-s; F(\xi(s)))\; ds \\
&\quad + \int_0^{t+\sigma} y(t+\sigma-s; G(\xi(s)))\; dM(s) \\
&= \pi_1\xi(t+\sigma) \quad \text{almost surely.}
\end{aligned}$$

The second equality holds since $y(r;\zeta) = 0$ for $-1 \leq r < 0$ if $\pi_2\zeta = 0$. We conclude that $\pi_2\xi$ is the segment process of $\pi_1\xi$.

Note that, by Fubini,

$$\int_0^t \int_0^s T(s-r)F(\xi(r))\; dr\; ds = \int_0^t \int_r^t T(s-r)F(\xi(r))\; ds\; dr, \quad t \geq 0,$$

and since $\int_r^t T(s-r)x\,ds \in \mathfrak{D}(A)$ for $x \in \mathcal{E}^2$, $0 \le r \le t$, (for any strongly continuous semigroup, see [32], Lemma II.1.3) we have, by closedness of A,

$$\int_0^t \int_0^s T(s-r)F(\xi(r))\,dr\,ds \in \mathfrak{D}(A), \quad t \ge 0, \tag{3.6}$$

Similarly, using stochastic Fubini (see [71], Theorem 8.14),

$$\int_0^t \int_0^s T(s-r)G(\xi(r))\,dM(r)\,ds \in \mathfrak{D}(A), \quad t \ge 0. \tag{3.7}$$

Now evaluate for $t \ge 0$,

$$\begin{aligned}
\pi_1 A \int_0^t \xi(s)\,ds &= \pi_1 A \int_0^t T(s) \begin{pmatrix} x \\ f \end{pmatrix} ds + \pi_1 A \int_0^t \int_0^s T(s-r)F(\xi(r))\,dr\,ds \\
&\quad + \pi_1 A \int_0^t \int_0^s T(s-r)G(\xi(r))\,dM(r)\,ds \\
&= \pi_1 A \int_0^t T(s) \begin{pmatrix} x \\ f \end{pmatrix} ds + \pi_1 \int_0^t A \int_r^t T(s-r)F(\xi(r))\,ds\,dr \\
&\quad + \pi_1 \int_0^t A \int_r^t T(s-r)G(\xi(r))\,ds\,dM(r) \\
&= \pi_1 (T(t) - I) \begin{pmatrix} x \\ f \end{pmatrix} + \pi_1 \int_0^t T(t-r)F(\xi(r)) - F(\xi(r))\,dr \\
&\quad + \pi_1 \int_0^t T(t-r)G(\xi(r)) - G(\xi(r))\,dM(r) \quad \text{almost surely.}
\end{aligned}$$

Note that all the terms are welldefined; in particular, A is evaluated only on elements in its domain by (3.6), (3.7).

Hence

$$\begin{aligned}
&x + \pi_1 A \int_0^t \xi(s)\,ds + \pi_1 \int_0^t F(\xi(s))\,ds + \pi_1 \int_0^t G(\xi(s))\,dM(s) \\
&= \pi_1 T(t) \begin{pmatrix} x \\ f \end{pmatrix} + \pi_1 \int_0^t T(t-r)F(\xi(r))\,dr + \pi_1 \int_0^t T(t-r)G(\xi(r))\,dM(r) \\
&= \pi_1 \xi(t) \quad \text{almost surely.}
\end{aligned}$$

We see that $\pi_1 \xi$ satisfies (3.4).

The converse statement is proved by a reversion of the above argument. □

Lemma 3.10. *There exists at most one solution $(Y(t))_{t \ge 0}$ to (3.4) satisfying*

$$\sup_{0 \le t \le T} \mathbb{E}|(Y(t), Y_t)|^2_{\mathcal{E}^2} < \infty \tag{3.8}$$

for all $T > 0$.

PROOF: Suppose there exist two solutions, $(Y(t))$ and $(Z(t))$, satisfying (3.8). Denote $\xi(t) := \begin{pmatrix} Y(t) \\ Y_t \end{pmatrix}$ and $\eta(t) := \begin{pmatrix} Z(t) \\ Z_t \end{pmatrix}$.

Then by a standard argument involving Gronwall's lemma, it may be shown that for $T > 0$

$$\sup_{0 \le t \le T} \mathbb{E}|\xi(t) - \eta(t)|^2_{\mathcal{E}^2} = 0.$$

□

Corollary 3.11. *Under the conditions of Hypothesis 3.7 there exists a unique solution to (3.5), or equivalently, to (3.4) by the identification* $\xi(t) = \begin{pmatrix} Y(t) \\ Y_t \end{pmatrix}$, $t \ge 0$, *satisfying (3.8).*

PROOF: This is a combination of Theorem 2.27, Theorem 3.9 and Lemma 3.10. □

3.3 Stochastic functional differential equation in finite dimensions

The most elementary example of the theory above is the case where $X = Z = \mathbb{R}^n$, so $\mathcal{E}^2 = \mathbb{R}^n \times L^2([-1,0];\mathbb{R}^n)$. In this case $B \in L(\mathbb{R}^n)$ is automatically bounded. Furthermore the injection of the noise is automatically Hilbert-Schmidt. A special case is the case for which Φ has the form

$$\Phi x := \sum_{i=1}^k B_i x(t - \theta_i) = \int_{-1}^0 d\eta \; x_t, \quad x \in W^{1,2}([-1,0];\mathbb{R}^n),$$

with $k \in \mathbb{N}$, $B_i \in L(\mathbb{R}^n)$ and $\theta_i \in [-1, 0]$ for $i = 1, \ldots, k$ and

$$\eta(\sigma) = \sum_{i=1}^k H(\sigma - \theta_i) B_i, \quad \sigma \in [-1, 0]$$

where H denotes the Heaviside step function.

We also assume that the noise assumes values in a finite dimensional space $\mathbb{R}^k$ so we allow as perturbations the mappings $\varphi : \mathcal{E}^2 \to \mathbb{R}^n$ and $\psi : \mathcal{E}^2 \to L(\mathbb{R}^k; \mathbb{R}^n)$. We thus arrive at the following typical example of a stochastic differential equation with delay:

$$dx(t) = \left[Bx(t) + \sum_{i=1}^k B_i x(t - h_i) + \varphi(x(t), x_t) \right] dt + \psi(x(t), x_t) \; dM(t). \quad (3.9)$$

3.3.1 Example: Population growth with random migration

Consider, as in Chapter 1, a population $(x(t))_{t\geq 0}$ evolving with constant birth rate $\beta > 0$ and constant death rate $\alpha > 0$. Let $r = 1$ indicate the development period of an individual. Suppose there is migration with random rate g which may depend on the size of the population. This leads to the stochastic differential equation with delay

$$dx(t) = [-\alpha x(t) + \beta x(t-r)]\ dt + g(x(t))\ dW(t), \quad t > 0.$$

For results on stability of this equation in the case of multiplicative noise, see Example 6.30.

3.4 Stochastic partial differential equations with delay

The following examples illustrate the two different cases of Hypothesis 3.7 (v).

3.4.1 Example: PDE with delay and noise

Consider the following reaction diffusion equation with delay and noise on a closed interval $\mathcal{O} \subset \mathbb{R}$:

$$\begin{cases} du(t,\xi) & \\ = \Big[\frac{\partial^2}{\partial \xi^2} u(t,\xi) + \int_0^1 \rho(s) u(t-s,\xi)\ ds & \\ \quad + \varphi(u(t),u_t)(\xi)\Big]\ dt + \Psi\ dM(t,\xi), & t \geq 0, \xi \in \mathcal{O}, \\ \frac{\partial u(t,\xi)}{\partial \nu} = 0, & t \geq 0, \xi \in \partial\mathcal{O} \\ u(t,\xi) = f(t,\xi), & t \in [-1,0], \xi \in \mathcal{O}, \\ u(0,\xi) = x(\xi), & \xi \in \mathcal{O}. \end{cases} \tag{3.10}$$

with

(i) delay distribution density $\rho \in L^\infty([0,1])$,

(ii) initial conditions $f \in L^2([-1,0]; L^2(\mathcal{O}))$ and $x \in L^2(\mathcal{O})$,

(iii) Lipschitz reaction term $\varphi : L^2(\mathcal{O}) \times L^2([-1,0]; L^2(\mathcal{O})) \to L^2(\mathcal{O})$ (possibly non-linear and/or non-local),

(iv) $(M(t))_{t\geq 0}$ a cylindrical martingale with RKHS $\mathcal{H} = L^2(\mathcal{O})$,

(v) $\Psi \in L(H)$.

Examples of cylindrical martingales with RKHS $L^2(\mathcal{O})$ are given in Section 2.4.

We employ the semigroup approach by setting $X = Z = L^2(\mathcal{O})$, as state space the Hilbert space $\mathcal{E}^2 = L^2(\mathcal{O}) \times L^2([-1,0]; L^2(\mathcal{O}))$, with A as defined in (3.2), with

$$B := \Delta, \quad \mathfrak{D}(B) = \left\{ v \in W^{1,2}(\mathcal{O}) : \Delta v \in L^2(\mathcal{O}) \text{ and } \frac{\partial v}{\partial \nu}(\xi) = 0, \quad \xi \in \partial\mathcal{O} \right\}$$

and

$$\Phi(w)(\xi) := \int_0^1 \rho(s) w(t-s,\xi)\, ds, \quad w \in L^2([-1,0]; L^2(\mathcal{O})).$$

Note that $\Phi \in L(L^2([-1,0]; L^2(\mathcal{O})))$. Furthermore let

$$F\left(\begin{pmatrix} v \\ w \end{pmatrix}\right) := \begin{pmatrix} \varphi(v,w) \\ 0 \end{pmatrix} \quad \text{and} \quad G(\cdot) := \begin{pmatrix} \Psi \\ 0 \end{pmatrix}, \quad \begin{pmatrix} v \\ w \end{pmatrix} \in \mathcal{E}^2.$$

Then the partial differential equation with noise and delay (3.10) is described by the abstract stochastic differential equation (3.5) in the state space $H = \mathcal{E}^2$. It may be verified that all conditions of Hypothesis 3.7 are satisfied. In particular it can be calculated explicitly that $t \mapsto ||S(t)||_{L_{\mathrm{HS}}(L^2(\mathcal{O}))}$ is square integrable on $[0,\infty)$, so that condition (v-b) of the Hypothesis holds. In more than one dimension this is not the case (see also [24], Example 5.7). We may conclude that there exists a unique solution to (3.5) in this context of this example, and hence there exists a unique solution to (3.10).

3.4.2 Example: Spreading and growing pollution with cleaning

Let $\mathcal{O} = [0,\pi]^d$, and let the noise and state spaces be defined as $U = L^2(\mathcal{O},\mu)$ and $H = L^2(\mathcal{O}, \mathrm{Leb})$, respectively. Suppose this region is polluted by random deposits at random locations with distribution μ on $\mathcal{O}$, modeled by the stochastic process $(Z(t))_{t\geq 0}$ of Section 2.4.2. The pollution spreads through diffusion at rate γ but cannot leave through the boundary. Furthermore cleaning takes place proportionally to the total amount of pollution present at rate κ. After one time step the pollution starts growing at rate η (think, for example, of batteries that start leaking).

The stochastic partial differential equation with delay corresponding to this description is

$$\begin{cases} du(t,\xi) & = (\gamma\Delta u(t,\xi) - \kappa u(t,\xi) + \eta u(t-1,\xi)\; dt \\ & \qquad +(\rho \star dZ(t))(\xi) \qquad\qquad t > 0, \xi \in \mathcal{O}, \\ \frac{\partial u(t,\xi)}{\partial \nu} & = 0, \qquad\qquad\qquad\qquad\qquad t > 0, \xi \in \partial\mathcal{O}, \\ u(0,\xi) & = x(\xi), \qquad\qquad\qquad\qquad \xi \in \mathcal{O}, \\ u(t,\xi) & = f(t,\xi), \qquad\qquad\qquad\quad -1 \leq t \leq 0, \xi \in \mathcal{O}. \end{cases} \tag{3.11}$$

Here $\kappa > 0$, $f \in L^2([-1,0]; L^2(\mathcal{O}))$ and $x \in L^2(\mathcal{O})$. The noise is convoluted with a mollifier $\rho \in L^2(\mathbb{R}^d)$ to ensure wellposedness of the equation, with convolution defined by

$$(\rho \star u)(\xi) := \int_{\mathcal{O}} \rho(\xi - \eta) u(\eta) \; d\mu(\eta), \quad \xi \in \mathcal{O}.$$

Note that the mapping

$$\Psi : U \to H, \quad u \mapsto \rho \star u$$

is Hilbert-Schmidt by the Hilbert-Schmidt kernel theorem (see [81], Proposition A.6.10 or [71], Proposition A.7) and the observation that the mapping

$$K(\xi, \eta) := \rho(\xi - \eta), \quad \xi, \eta \in \mathcal{O}$$

satisfies $K \in L^2(\mathcal{O} \times \mathcal{O}, \mathrm{Leb} \otimes \mu)$. Indeed,

$$\begin{aligned} ||\Psi||^2_{L_{\mathrm{HS}}(U;H)} &= \int_{\mathcal{O}} \int_{\mathcal{O}} K(\xi,\eta)^2 \; d\xi \; d\mu(\eta) \leq \int_{\mathcal{O}} \int_{\mathbb{R}^d} \rho(\xi - \eta)^2 \; d\xi \; d\mu(\eta) \\ &= \int_{\mathcal{O}} \int_{\mathbb{R}^d} \rho(\zeta)^2 \; d\zeta \; d\mu(\eta) = ||\rho||^2_{L^2(\mathbb{R}^d)} \mu(\mathcal{O}) = ||\rho||^2_{L^2(\mathbb{R}^d)}, \end{aligned}$$

where we applied the translation $\zeta = \xi - \eta$.

In a similar way as with the previous example we may treat this problem as a stochastic evolution equation in $L^2(\mathcal{O}) \times L^2([-1,0]; L^2(\mathcal{O}))$ but now we do not need the delay semigroup to be of bounded Hilbert-Schmidt norm since the injection of the noise in $L^2(\mathcal{O})$ is Hilbert-Schmidt. As generator of the delay semigroup we have

$$A = \begin{bmatrix} B & \Phi \\ 0 & \frac{d}{d\sigma} \end{bmatrix} \tag{3.12}$$

with

$$B = \gamma \Delta - \kappa I \quad \text{and} \quad \Phi z = \eta z(-1), \quad z \in W^{1,2}([-1,0]; L^2(\mathcal{O})).$$

Recall (see Section 2.4.2) that

$$Z(t) = \sum_{\tau_k \leq t} \sigma_k \delta_{\xi_k}, \quad t \geq 0$$

The meaning of the stochastic part $\rho \star dZ(t)$ is made clear by the following formal manipulation:

$$(\rho \star Z(t))(\xi) = \int_{\mathcal{O}} \rho(\xi - \eta) \sum_{\tau_k \leq t} \sigma_k \delta_{\xi_k}(\eta) \; d\mu(\eta) = \sum_{\tau_k \leq t} \sigma_k \rho(\xi - \xi_k), \quad t \geq 0, \xi \in \mathcal{O}.$$

Let

$$\psi(\xi) := \frac{1}{t} \mathbb{E}\left[\rho \star Z(t)\right] = \lambda \int_0^\infty \sigma \; d\nu(\sigma) \int_{\mathcal{O}} \rho(\xi - \eta) \; d\mu(\eta), \quad t \geq 0,$$

again using notation of Section 2.4.2. If we write $N(t) := \rho \star Z(t) - \mathbb{E}[\rho \star Z(t)]$, then N is a square integrable martingale and we may rewrite the stochastic equation of (3.11) as

$$du(t,\xi) = (\gamma \Delta u(t,\xi) - \kappa u(t,\xi) + \eta u(t-1,\xi) + \psi(\xi)\, dt + dN(t) \quad t > 0, \xi \in \mathcal{O}. \tag{3.13}$$

See Section 3.6.1 for results on dissipativeness of the semigroup corresponding to this stochastic evolution and Section 4.2.3 and Section 4.3.1 for the existence (and uniqueness) of an invariant measure for this stochastic evolution.

3.5 Eventual compactness of the delay semigroup

Definition 3.12. *A strongly continuous semigroup $(T(t))_{t\geq 0}$ with infinitesemal generator A is called* eventually compact *if there exists a $t_0 > 0$ such that $T(t_0)$ is a compact operator.*

Under the condition of compactness of $(S(t))_{t\geq 0}$, we can show eventual compactness for the delay semigroup $(T(t))_{t\geq 0}$ as described in Section 3.1.

To show eventual compactness of the delay semigroup we proceed as in [32], Section VI.6. We make use of the following variant of the Arzelà-Ascoli theorem.

Definition 3.13. *A subset Φ of $C(X;Y)$ is* pointwise relatively compact *if and only if for all $x \in X$, the set $\{f(x) : f \in \Phi\}$ is relatively compact in Y.*

Theorem 3.14 (vector valued Arzelà-Ascoli, [66], Theorem 47.1)**.** *Let X be a compact Hausdorff space and Y a metric space. Then a subset Φ of $C(X;Y)$ is relatively compact if and only if it is equicontinuous and pointwise relatively compact.*

Lemma 3.15. *Suppose $(S(t))_{t\geq 0}$ is immediately compact. Then $R(\lambda, A)T(1)$ is compact for all $\lambda \in \rho(A)$.*

Proof: According to [7], Proposition 3.19, we have the following expression for the resolvent $R(\lambda, A)$:

$$R(\lambda, A) = \begin{bmatrix} R(\lambda, B + \Phi_\lambda) & R(\lambda, B + \Phi_\lambda)\Phi R(\lambda, A_0) \\ \epsilon_\lambda R(\lambda, B + \Phi_\lambda) & [\epsilon_\lambda R(\lambda, B + \Phi_\lambda)\Phi + I]R(\lambda, A_0) \end{bmatrix}, \quad \lambda \in \rho(A), \tag{3.14}$$

where, for $\lambda \in \mathbb{C}$, $\Phi_\lambda \in L(X)$ is given by

$$\Phi_\lambda x := \Phi(e^{\lambda \cdot} x), \quad x \in X,$$

ϵ_λ is the function

$$\epsilon_\lambda(s) := \mathrm{e}^{\lambda s}, \quad s \in [-1, 0].$$

and A_0 is the generator of the nilpotent left-shift semigroup on $L^p([-1,0]; X)$.

Let

$$\pi_1 : X \times L^p([-1,0];Z) \to X$$

and

$$\pi_2 : X \times L^p([-1,0];Z) \to L^p([-1,0];Z)$$

denote the canonical projections on X and $L^p([-1,0];Z)$, respectively.

Lemma 4.5 and Lemma 4.9 in [7] state that the operator $R(\lambda, B+\Phi_\lambda)$ is compact for all $\lambda \in \rho(A)$. Therefore, using (3.14)

$$\pi_1 R(\lambda, A)T(1) = \begin{bmatrix} R(\lambda, B+\Phi_\lambda) & R(\lambda, B+\Phi_\lambda)\Phi R(\lambda, A_0) \end{bmatrix} T(1) \tag{3.15}$$

is compact.

We can therefore restrict our attention to

$$\pi_2 R(\lambda, A)T(1) : X \times L^p([-1,0];Z) \to L^p([-1,0];Z).$$

Denote $\varphi := \begin{pmatrix} x \\ f \end{pmatrix}$ where $x \in X$ and $f \in L^p([-1,0];Z)$. Note that

$$\frac{d}{d\sigma}\pi_2 R(\lambda, A)T(1)\varphi = \pi_2 AR(\lambda, A)T(1)\varphi,$$

Hence, using Hölder, there exists some constant $M > 0$ such that for all $t_0, t_1 \in [-1,0]$,

$$\begin{aligned}
||\pi_2 R(\lambda, A)T(1)\varphi\,[(t_1) - (t_0)]||_Z &= \left|\left| \int_{t_0}^{t_1} \left[\frac{d}{d\sigma}\pi_2 R(\lambda, A)T(1)\varphi\right](\sigma)\; d\sigma \right|\right|_Z \\
&= \left|\left| \int_{t_0}^{t_1} \left[\pi_2 AR(\lambda, A)T(1)\varphi\right](\sigma)\; d\sigma \right|\right|_Z \\
&\le \int_{t_0}^{t_1} ||[\pi_2 AR(\lambda, A)T(1)\varphi](\sigma)||_Z \; d\sigma \\
&\le |t_1 - t_0|^{1/q}\, ||\pi_2 AR(\lambda, A)T(1)\varphi||_{L^p([-1,0];Z)} \\
&\le M|t_1 - t_0|^{1/q}\, ||\varphi||_{\mathcal{E}^p}\,.
\end{aligned}$$

Here $q > 1$ is such that $\frac{1}{q} + \frac{1}{p} = 1$.

So

$$\mathcal{C} := \{\pi_2 R(\lambda, A)T(1)\varphi : \varphi \in \mathcal{E}^p, ||\varphi||_{\mathcal{E}^p} \le 1\} \subset C([-1,0];Z)$$

is equicontinuous.

Furthermore note that

$$\begin{aligned}
&[\pi_2 R(\lambda, A)T(1)\varphi](\sigma) \\
&= [\pi_2 T(1)R(\lambda, A)\varphi](\sigma) && \text{(commutativity)} \\
&= [\pi_2 T(1+\sigma)R(\lambda, A)\varphi](0) && \text{(translation property)} \\
&= [\pi_2 R(\lambda, A)T(1+\sigma)\varphi](0) && \text{(commutativity)} \\
&= \pi_1 R(\lambda, A)T(1+\sigma)\varphi && (R(\lambda, A) \text{ maps into } \mathfrak{D}(A)).
\end{aligned}$$

Again using (3.14) and the fact that $R(\lambda, B + \Phi_\lambda)$ is compact, we find that $\mathcal{C}$ is pointwise relatively compact. By the vector-valued Arzelà-Ascoli theorem, Theorem 3.14, we find that $\mathcal{C}$ is relatively compact in $C([-1, 0]; Z)$ and hence relatively compact in $L^p([-1, 0]; Z)$.

From this we conclude that $\pi_2 R(\lambda, A)T(1)$ is compact and combining this with (3.15), $R(\lambda, A)T(1)$ is compact. $\square$

We may now conclude that $(T(t))_{t\geq 0}$ is eventually compact:

Theorem 3.16. *Suppose Hypothesis 3.3 holds. Furthermore suppose $(S(t))_{t\geq 0}$ is immediately compact. Then $(T(t))_{t\geq 0}$ is compact for all $t > 1$.*

Proof: By [32], Lemma II.4.28, it is sufficient to show that $(T(t))$ is eventually norm continuous for $t > 1$, and that $R(\lambda, A)T(1)$ is compact for some $\lambda \in \rho(A)$.

Now by [7], Lemma 4.5, $(T(t))$ is norm continuous for $t > 1$ (using that $(S(t))$ is immediately compact and hence immediately norm continuous). Furthermore by Lemma 3.15, $R(\lambda, A)T(1)$ is compact for all $\lambda \in \rho(A)$. $\square$

3.6 Dissipativeness of the delay semigroup

A convenient property that some generators of strongly continuous semigroups possess is that of dissipativity.

Definition 3.17. *Suppose $A : \mathfrak{D}(A) \to H$ is a linear operator. Then A is said to be* dissipative *if $\langle Ax, x\rangle \leq 0$ for all $x \in \mathfrak{D}(A)$.*

A strongly continuous semigroup $(T(t))_{t\geq 0}$ is said to be a contraction semigroup *if $||T(t))|| \leq 1$ for all $t \geq 0$.*

It is well-known that a strongly continuous semigroup is a contraction semigroup if and only if the infinitesemal generator of the semigroup is dissipative.

The fact that a semigroup is a contraction semigroup is important in its own right. Moreover, in the context of stochastic evolutions dissipativity of the generator can be used in many places, see for example Theorem 2.32, which establishes the cadlag property of paths, Theorem 4.11 concerning the existence and uniqueness of invariant measures, and Theorems 6.15 and 6.26 which establish the stability of stochastic evolutions.

Recall the notation of Section 3.1. In this section we assume $X = Z$ to be a Hilbert space and $p = 2$.

Theorem 3.18. *Suppose Φ has the form*

$$\Phi f = \int_{-1}^{0} d\zeta(s)\ f(s),$$

with $\zeta : [-1, 0] \to L(X)$ of bounded variation. Suppose furthermore that $B - \lambda I$ is dissipative. If there exists $\mu > \lambda$ such that

$$(\lambda - \mu)^2 > (|\zeta|(0) - |\zeta|(-1)) \int_{-1}^{0} \mathrm{e}^{2\mu r}\ d|\zeta|(r),$$

then there exists an equivalent inner product on $X \times L^2([-1, 0]; X)$ such that $A - \mu I$ is dissipative with respect to this inner product.

Proof: Introduce a quadratic form on $X \times L^2([-1, 0]; X)$ by

$$\left(\begin{pmatrix} c \\ f \end{pmatrix}, \begin{pmatrix} d \\ g \end{pmatrix} \right) := \langle c, d\rangle + \int_{-1}^{0} \tau(s)\langle f(s), g(s)\rangle\ ds.$$

We will make this into the new inner product satisfying the mentioned requirements. First we determine τ such that $\langle (A - \mu I)x, x\rangle \leq 0$ for all $x \in \mathfrak{D}(A)$. Then we will check that $\langle \cdot, \cdot \rangle$ is an equivalent inner product on $X \times L^2([-1, 0]; X)$.

We have for $x \in \mathfrak{D}(A)$,

$$\begin{aligned}
&((A - \mu I)x, x) \\
&= \langle Bx(0), x(0)\rangle + \left\langle \int_{-1}^{0} d\zeta(s)x(s), x(0) \right\rangle \\
&\quad + \int_{-1}^{0} \tau(s)\langle \dot{x}(s), x(s)\rangle\ ds - \mu|x(0)|^2 - \mu \int_{-1}^{0} \tau(s)|x(s)|^2\ ds \\
&\leq (\lambda - \mu)|x(0)|^2 + \int_{-1}^{0} |x(0)||x(s)|\ d|\zeta|(s) + \int_{-1}^{0} \tau(s) \left(\tfrac{1}{2}\frac{d}{ds}|x(s)|^2 - \mu|x(s)|^2 \right)\ ds \\
&= (\lambda - \mu)|x(0)|^2 + \int_{-1}^{0} |x(0)||x(s)|\ d|\zeta|(s) - \int_{-1}^{0} \tfrac{1}{2}\dot{\tau}(s)|x(s)|^2\ ds \\
&\quad + \tfrac{1}{2}\left\{\tau(0)|x(0)|^2 - \tau(-1)|x(-1)|^2\right\} - \int_{-1}^{0} \mu\tau(s)|x(s)|^2\ ds \\
&\leq (\lambda - \mu + \tfrac{1}{2}\gamma)|x(0)|^2 + \int_{-1}^{0} |x(0)||x(s)|\ d|\zeta|(s) + \int_{-1}^{0} \left(-\mu\tau(s) - \tfrac{1}{2}\dot{\tau}(s)\right)|x(s)|^2\ ds,
\end{aligned}$$

where we denote $\gamma := \tau(0)$. Since $\dot\tau$ may not be defined, the last steps have to be interpreted formally.

In order to be able to compare the two integrals, we demand (formally) that

$$\left(-\mu\tau(s) - \tfrac{1}{2}\dot\tau(s)\right)\ ds = \rho(s) d|\zeta|(s), \tag{3.16}$$

with $\rho : [-1, 0] \to \mathbb{R}$ some, for the moment unspecified, function.

Under this assumption, we obtain as sufficient condition for dissipativity

$$(-\mu + \lambda + \tfrac{1}{2}\gamma)|x(0)|^2 + \int_{-1}^0 |x(0)||x(s)| + \rho(s)|x(s)|^2\ d|\zeta|(s) \le 0.$$

Dividing the lefthand side by $|x(0)|^2$, we obtain

$$\begin{aligned}
&(-\mu + \lambda + \tfrac{1}{2}\gamma) + \int_{-1}^0 \frac{|x(s)|}{|x(0)|} + \rho(s)\left(\frac{|x(s)|}{|x(0)|}\right)^2\ d|\zeta(s)| \\
&= \int_{-1}^0 \frac{-\mu + \lambda + \frac{1}{2}\gamma}{|\zeta|(0) - |\zeta|(-1)} + \frac{|x(s)|}{|x(0)|} + \rho(s)\left(\frac{|x(s)|}{|x(0)|}\right)^2\ d|\zeta(s)| \le 0.
\end{aligned}$$

Note that the integrand is a polynomial in $\frac{|x(s)|}{|x(0)|}$, which is at most zero if and only if

$$\rho(s) \le 0, \quad \text{a.a.} \quad s \in [-1, 0],$$

and

$$1 - 4\rho(s)\frac{-\mu + \lambda + \frac{1}{2}\gamma}{|\zeta|(0) - |\zeta|(-1)} \le 0.$$

We can obtain equality by putting $\rho(s) \equiv \rho$, with

$$\rho := \frac{|\zeta|(0) - |\zeta|(-1)}{4\left(-\mu + \lambda + \frac{1}{2}\gamma\right)}.$$

Since necessarily $\rho \le 0$, we require

$$-\mu + \lambda + \tfrac{1}{2}\gamma < 0.$$

It is a straightforward exercise in partial integration to show that the expression

$$\tau(s) := \mathrm{e}^{-2\mu s}\left[\gamma - \frac{|\zeta|(0) - |\zeta|(-1)}{-2\lambda + 2\mu - \gamma}\int_s^0 \mathrm{e}^{2\mu r}\ d|\zeta|(r)\right]$$

formally solves (3.16), and actually satisfies the required

$$\begin{aligned}
&\int_{-1}^0 \tau(s)\left(\tfrac{1}{2}\frac{d}{ds}|x(s)|^2 - \mu|x(s)|^2\right)\ ds \\
&= \int_{-1}^0 \rho|x(s)|^2\ d|\zeta|(s) + \tfrac{1}{2}\gamma|x(0)|^2 - \tfrac{1}{2}\tau(-1)|x(-1)|^2.
\end{aligned}$$

For $\langle\cdot,\cdot\rangle$ to define an equivalent inner product on $X \times L^2([-1,0];X)$, we require that there exist constants $c_1, c_2 > 0$ such that $c_1 \leq |\tau(s)| \leq c_2$ for almost all $s \in [-1,0]$. In particular we demand that $\tau(s) > 0$, s-a.s.

Note that $\tau(s) > 0$ if and only if

$$\gamma - \frac{|\zeta|(0) - |\zeta|(-1)}{-2\lambda + 2\mu - \gamma} \int_s^0 \mathrm{e}^{2\mu r}\, d|\zeta|(r) > 0, \quad \text{a.a. } s \in [-1,0],$$

that is

$$\gamma - \frac{|\zeta|(0) - |\zeta|(-1)}{-2\lambda + 2\mu - \gamma} \int_{-1}^0 \mathrm{e}^{2\mu r}\, d|\zeta|(r) > 0$$

or equivalently

$$\gamma(-2\lambda + 2\mu - \gamma) > (|\zeta|(0) - |\zeta|(-1)) \int_{-1}^0 \mathrm{e}^{2\mu r}\, d|\zeta|(r).$$

We are free to choose γ in $(0, 2(-\lambda + \mu))$, so we pick the optimal $\gamma = -\lambda + \mu$. The requirement

$$(-\lambda + \mu)^2 > (|\zeta|(0) - |\zeta|(-1)) \int_{-1}^0 \mathrm{e}^{-2\mu r}\, d|\zeta|(r)$$

remains. If it is satisfied, then it is easy to check that τ is bounded from below and above and therefore does indeed define an equivalent inner product. $\square$

Corollary 3.19. *Suppose Φ has the form*

$$\Phi f = \int_{-1}^0 d\zeta(s)\ f(s),$$

with $\zeta : [-1,0] \to L(X)$ of bounded variation. Suppose furthermore that $B - \lambda I$ is dissipative. If

$$\lambda < -\int_{-1}^0 d|\zeta|,$$

then there exists an equivalent inner product on $\mathcal{E}^2$ such that A is dissipative.

PROOF: This follows immediately by noting that $\lambda^2 > (|\zeta|(0) - |\zeta|(-1))^2$ so that the condition of Theorem 3.18 is satisfied with $\mu = 0$. $\square$

In particular this gives us a stability result for the delay semigroup. The estimate provided is 'sharp' in the sense that it reproduces the stability result [7], Corollary 5.9.

Theorem 3.20. *Suppose B is the generator of a generalized contraction semigroup. Then there exists an equivalent inner product on $\mathcal{E}^2$ such that A is the generator of a generalized contraction semigroup.*

PROOF: Denote the semigroup generated by B by $(S(t))_{t\geq 0}$ and suppose that

$$||S(t)|| \leq e^{\lambda t}, \quad t \geq 0.$$

Let $\nu > \max\left(0, \lambda + \int_{-1}^{0} d|\zeta|\right)$. Define $\widetilde{B} := B - \nu I$ and $\widetilde{\lambda} := \lambda - \nu$. It may be verified that the conditions of Corollary 3.19 are satisfied for $\widetilde{B}$ and $\widetilde{\lambda}$, so that an equivalent inner product $(\cdot, \cdot)$ on $\mathcal{E}^2$ exists such that the delay semigroup generated by

$$\widetilde{A} := \begin{bmatrix} B - \nu I & \Phi \\ 0 & \frac{d}{d\sigma} \end{bmatrix}$$

is dissipative. Now for $x \in \mathfrak{D}(A)$

$$((A - \nu I)x, x) = (\widetilde{A}x, x) - \nu \int_{-1}^{0} \tau(\sigma) x(\sigma)^2 \, d\sigma \leq (\widetilde{A}x, x) \leq 0.$$

□

Note furthermore that, if A is of the form (3.2) with $B - \lambda I$ dissipative for some $\lambda \in \mathbb{R}$, we can always perturb A by a bounded operator of the form $\begin{pmatrix} -cI & 0 \\ 0 & 0 \end{pmatrix}$ to obtain the generator of a new delay semigroup. If we choose $c > 0$ large enough, by Corollary 3.19 we may change the inner product to obtain a dissipative generator.

In an entirely analogous way as for Theorem 3.18 we can prove the following slightly stronger result in case Φ has a density function.

Proposition 3.21. *Suppose Φ is of the form*

$$\Phi f = \int_{-1}^{0} \zeta(\sigma) f(\sigma) \, d\sigma, \quad f \in L^2([-1, 0]; X),$$

with $\zeta \in L^2([-1, 0]; L(X))$. Suppose furthermore $B - \lambda I$ is dissipative. If there exists $\mu > \lambda$ such that

$$(\lambda - \mu)^2 > \int_{-1}^{0} e^{2\mu\rho} ||\zeta(\rho)||^2 \, d\rho.$$

then there exists an equivalent inner product on $\mathcal{E}_2$ such that $A - \mu I$ is dissipative with respect to this inner product.

The following is a first attempt at establishing conditions on more general generators A of strongly continuous semigroups such that there exists an equivalent inner product such that A is dissipative.

To this end we need the notion of exact observability.

Definition 3.22. *Let A be the generator of a strongly continuous semigroup $(T(t))_{t\geq 0}$ in a Hilbert space H. Let K be a Hilbert space and $C \in L(H; K)$. Then (A, C) is said to be* exactly observable *in time $\tau > 0$ if the mapping $x \mapsto C^\tau x := CT(\cdot)x : H \to L^2([0, \tau]; K)$ is injective and its inverse is bounded on the range of C^τ.*

Observability is a concept dual to controllability: (A, C) is exactly observable if and only if (A, C^*) is exactly controllable (see [20], Section 4.1).

Proposition 3.23. *Suppose A generates an asymptotically stable strongly continuous semigroup $(T(t))_{t\geq 0}$ in H. If there exists a Hilbert space K, a bounded linear mapping $C \in L(H; K)$ and $\tau > 0$ such that (A, C) is exactly observable in time τ, then there exists an equivalent inner product on H such that A is dissipative with respect to this inner product.*

PROOF: For this C, by the characterization of exactly observable systems ([20], Corollary 4.1.14) we have

$$Q := \int_0^\infty T^*(t)C^*CT(t)\, dt \geq \int_0^\tau T^*(t)C^*CT(t)\, dt \geq \gamma I$$

for some $\gamma > 0$. Furthermore by Lyapunov theory ([20], Theorem 4.1.23) we have that $Q \in L(H)$ and

$$2\langle QAx, x\rangle = -|Cx|^2 \leq 0, \quad x \in \mathfrak{D}(A).$$

□

3.6.1 Example: Environmental pollution

We return to the example of Section 3.4.2, where

$$A = \begin{bmatrix} B & \Phi \\ 0 & \frac{d}{d\sigma} \end{bmatrix}$$

with

$$B = \gamma\Delta - \kappa I \quad \text{and} \quad \Phi z = \eta z(-1), \quad z \in W^{1,2}([-1,0]; L^2(\mathcal{O}))$$

and where γ, κ and η are larger than zero.

Note that $B + \kappa I$ is dissipative, and therefore, by Corollary 3.19, the state space $L^2(\mathcal{O}) \times L^2([-1,0]; L^2(\mathcal{O}))$ may be equipped with a new inner product such that the generator of the delay semigroup is dissipative, as long as $|\eta| < \kappa$.

By tracing the proof of Theorem 3.18 this inner product can be seen to have the form

$$\left(\begin{pmatrix} c \\ x \end{pmatrix}, \begin{pmatrix} d \\ y \end{pmatrix}\right) = \langle c, d\rangle_{L^2(\mathcal{O})} + \kappa \int_{-1}^0 \langle x(s), y(s)\rangle_{L^2(\mathcal{O})}\, ds.$$

3.7 Notes and remarks

Throughout this chapter we assumed for simplicity that the maximum amount of delay equals one timestep but of course all results of this section can be extended to more general maximum delay amounts.

The proof of dissipativeness of the delay semigroup under an equivalent inner product is a generalisation (with some loss of strength) of a result of [25], Section 10.3. Based on the contents of Section 3.6, a short paper is being prepared for publication ([8]).

[63] is an introduction to stochastic delay differential equations. For results on stochastic delay differential in the space of continuous functions, see [89] and [90].

CHAPTER

FOUR

Existence of an invariant measure

We consider here infinite-dimensional diffusions in a Hilbert space H described by the differential equation

$$\begin{cases} dX(t) = [AX(t) + F(X(t))]\, dt + G(X(t))\, dM(t), & t \geq 0, \\ X(0) = x, \end{cases} \tag{4.1}$$

with A the generator of a strongly continuous semigroup, F and G Lipschitz functions and M a cylindrical martingale of stationary covariance, see Section 2.5.

For many choices of A, F and G it is impossible to obtain the exact solution of such an equation. Therefore it is important to establish qualititative properties of the solution on the basis of information on A, F, G and M.

One of these qualitative properties is the existence of an *invariant probability measure*: under what conditions does a measure μ on H exist such that if the initial condition x has distribution μ, we have that $X(t)$ has distribution μ for all $t \geq 0$.

Often a compactness argument (Krylov-Bogoliubov, see e.g. [25], Theorem 3.1.1) is used to establish the existence of an invariant measure. In finite dimensions it suffices to show that the solutions of (4.1) are bounded in probability. In infinite dimensions, due to the absence of local compactness, we need to exploit compactness properties of the solutions of the stochastic differential equation.

It has been shown [22] that a suitable criterion is that A generates a compact semigroup. Together with solutions bounded in probability this suffices to prove the existence of an invariant measure. Another approach is taken in [85], based on hyperbolicity of the driving semigroup and small Lipschitz coefficients of the perturbations.

The result obtained for compact semigroups leads immediately to the question whether eventual compactness of the semigroup can be used to establish existence of an invariant measure. This is an interesting question because, for example, delay differential equations, when put in an infinite-dimensional framework, possess this property (see Section 3.5). Also in the theory of structured population equations eventually compact semigroups appear (see e.g. [31] and [32], Section VI.1). In [25] it is conjectured that eventual compactness should be a sufficient criterion for the existence of an invariant measure.

In Section 4.1 we show that eventual compactness of the semigroup, together with compact factorizations of the perturbations F and G, can indeed be used to establish the existence of an invariant measure. The result is applied to some examples in Section 4.2, among which a stochastic functional differential equation and the currently active (see e.g. [53], [54]) field of reaction-diffusion equations perturbed by delayed feedback and noise (Section 4.2). In Section 4.3 we mention a result concerning the existence and uniqueness of invariant measures for dissipative systems.

4.1 Main result

Recall that a strongly continuous semigroup $(T(t))_{t\geq 0}$ with infinitesemal generator A is called eventually compact if there exists a $t_0 > 0$ such that $T(t_0)$ is a compact operator.

Throughout this section, we will assume that the following hypothesis holds:

Hypothesis 4.1. *(i) H and $\mathcal{H}$ are separable Hilbert spaces and E_1 and E_2 are Banach spaces;*

(ii) A is the generator of a strongly continuous, eventually compact semigroup $(S(t))$ on H; we assume without loss of generality that $S(t)$ is compact for $t \geq 1$;

(iii) $F : H \to H$ is globally Lipschitz and admits a factorization $F = C_1 \circ \Phi$, where $C_1 \in L(E_1; H)$ is compact and $\Phi : H \to E_1$ satisfies $|\Phi(x)| \leq K(1 + |x|)$ for all $x \in H$ and some constant $K > 0$;

(iv) $G : H \to L_{\mathrm{HS}}(\mathcal{H}; H)$ is globally Lipschitz and admits a factorization $G(x) = C_2\Psi(x)$, $x \in H$, where $C_2 \in L(E_2; H)$ is compact, and $\Psi : H \to L_{\mathrm{HS}}(\mathcal{H}; E_2)$ satisfies $|\Psi(x)|_{L_{\mathrm{HS}}(\mathcal{H};H)} \leq K(1 + |x|)$ for all $x \in H$ and some constant $K > 0$;

(v) M is a cylindrical martingale with RKHS $\mathcal{H}$;

(vi) For $x \in H$, $(X(t, x))_{t\geq 0, x\in H}$ is the unique mild solution (which exists by Theorem 2.27) of the stochastic differential equation

$$\begin{cases} dX(t) = (AX(t) + F(X))\ dt + G(X)\ dM(t), & t \geq 0, \\ X(0) = x, & x \in H \end{cases}$$

(vii) For all $x \in H$ and $\varepsilon > 0$, there exists $R > 0$ such that for all $T \geq 1$,

$$\frac{1}{T}\int_0^T \mathbb{P}(|X(t,x)| \geq R)\; dt < \varepsilon.$$

Remark 4.2. Note that (vii) of Hypothesis 4.1 is trivially satisfied if solutions $(X(t,x))$ are bounded in probability.

Under these assumptions, we can establish the existence of an invariant measure.

Theorem 4.3. *Suppose Hypothesis 4.1 is satisfied. Then there exists an invariant measure for $(X(t,x))_{t\geq 0}$.*

To prove the theorem we need a couple of lemmas.

Lemma 4.4. *Let $(T(t))$ be a strongly continuous semigroup acting on E_1, $C \in L(E_2; E_1)$ and suppose that $T(t)C$ is compact for all $t > 0$.*

Then

$$V_\gamma = \{T(t)Ck : t \in [\gamma, 1], k \in E_2, |k| \leq 1\}$$

is relatively compact for all $\gamma > 0$.

Moreover, if C is compact, then V_γ is relatively compact for all $\gamma \geq 0$.

Proof: Assume $T(t)C$ is compact for all $t > 0$. We will show that V_γ is relatively compact for $\gamma > 0$. Indeed, let (v_n) be a sequence in V_γ. There exist sequences $(t_n) \subset [\gamma, 1]$, $(x_n) \subset E_2$, with $|x_n| \leq 1$, $n \in \mathbb{N}$, such that

$$v_n = T(t_n)Cx_n = T(t_n - \gamma)T(\gamma)Cx_n, \quad n \in \mathbb{N}.$$

Since $T(\gamma)C$ is compact, there exists a subsequence (x_{n_k}) of (x_n) such that $T(\gamma)Cx_{n_k} \to y$ for some $y \in E_1$, $|y| \leq ||T(\gamma)C||$. Write $s_n := t_n - \gamma \in [0, 1-\gamma]$, $n \in \mathbb{N}$. Since $[0, 1-\gamma]$ is compact, by strong continuity of $(T(t))$, there exists a further subsequence $(s_{n_{k_l}})$ of (s_{n_k}) such that $T(s_{n_{k_l}})y \to z$ with $z \in T([0, 1-\gamma])y$.

Now

$$\begin{aligned}
|T(t_{n_{k_l}})Cx_{n_{k_l}} - z| &\leq |T(s_{n_{k_l}})T(\gamma)Cx_{n_{k_l}} - T(s_{n_{k_l}})y| + |T(s_{n_{k_l}})y - z| \\
&\leq ||T(s_{n_{k_l}})|||T(\gamma)Cx_{n_{k_l}} - y| + |T(s_{n_{k_l}})y - z| \\
&\leq m|T(\gamma)Cx_{n_{k_l}} - y| + |T(s_{n_{k_l}})y - z| \to 0
\end{aligned}$$

as $l \to \infty$. Here $m = \sup_{t\in[0,1-\gamma]} ||T(t)||$.

If $C = T(0)C$ is compact, then the proof above can be repeated for the case $\gamma = 0$. ☐

Lemma 4.5. *Let $(T(t))$ be a strongly continuous semigroup acting on E_1, $C \in L(E_2; E_1)$ and suppose that $T(t)C$ is compact for all $t > 0$. Let*

$$Gf := \int_0^1 T(1-s)Cf(s)\ ds, \quad f \in L^p([0,1]; E_2),$$

with $p > 1$. Then $G \in L(L^p([0,1]; E_2); E_1)$ is compact.

Proof: Define for $\gamma > 0$ the operators $G_\gamma \in L(L^p([0,1]; E_2); E_1)$ by

$$G_\gamma f = \int_0^{1-\gamma} T(1-s)Cf(s)\ ds, \quad f \in L^p([0,1]; E_2).$$

By Lemma 4.4, V_γ is relatively compact for $\gamma > 0$ and therefore its closed convex hull K_γ is compact ([76], Theorem 3.25).

Let $f \in L^p([0,1]; E_2)$ and define a positive measure on $[0,1]$ by

$$\mu_f(ds) := |f(s)|\ ds.$$

Note that μ_f is a finite measure since, by Jensen,

$$\mu_f([0,1]) = \int_0^1 |f(s)|\ ds \le \left(\int_0^1 |f(s)|^p\ ds\right)^{\frac{1}{p}}.$$

Now

$$G_\gamma f = \int_0^{1-\gamma} T(1-s)C\frac{f(s)}{|f(s)|}\ \mu_f(ds),$$

is an integral over a positive, finite measure with the integrand assuming values in the convex set K_γ, so

$$G_\gamma f \in \mu_f([0,1])K_\gamma = ||f||_{L^p([0,1];E_2)}K_\gamma.$$

So G_γ is compact for $\gamma > 0$.

Finally, it is straightforward, using Hölder and that $p > 1$, to show that $G_\gamma \to G$ in operator norm as $\gamma \downarrow 0$, so G is compact. $\square$

We will also need the following lemma from [13].

Lemma 4.6. *Let H be a separable Hilbert space. Let $K \subset H$ be compact. Then there exists a compact, self-adjoint, strictly positive definite operator $T \in L(H)$ such that*

$$K \subset \{Tx : |x| \le 1\}.$$

The proof in [13], Example 3.8.13(ii) is rather short for such an interesting lemma. We present a full proof here.

Proof: Assume for simplicity that we deal with ℓ^2 and that there exists a non-zero $x \in K$.

Claim:

$$\lim_{n\to\infty} \sup_{x\in K} \sum_{i=n}^{\infty} x_i^2 = 0.$$

Proof of claim: Suppose there exists a $\delta > 0$ such that for all $n \in \mathbb{N}$, there exists $x^n \in K$ such that

$$\sum_{i=n}^{\infty} (x_i^n)^2 \geq \delta.$$

Now for fixed n pick such an x^n and $m \in \mathbb{N}$ such that

$$\sum_{i=m}^{\infty} (x_i^n)^2 < \delta/2.$$

Furthermore pick x^m such that

$$\sum_{i=m}^{\infty} (x_i^m)^2 \geq \delta.$$

Then

$$||x^n - x^m||_{\ell^2}^2 \geq ||(x^n - x^m)\mathbb{1}_{\{m,m+1,\dots\}}||_{\ell^2}^2 > \delta/2.$$

So the sequence (x^n) does not have a Cauchy subsequence. Hence K is not compact, which proves the claim. $\diamond$

So we can find an increasing sequence $(N_n)_{n=1}^{\infty}$ such that

$$\sum_{i=N_n}^{\infty} x_i^2 \leq 4^{-n} \quad \text{for all } x \in K.$$

Let $t_i > 0$, $t_i^2 := 2^{-n+1}$ for $N_n \leq i < N_{n+1}$, $n \in \mathbb{N}$, and $t_i^2 := 2\sup_{x\in K} ||x||_{\ell^2}^2$ for $1 \leq i < N_1$. Define $T \in L(H)$ by $(Tx)_i := t_i x_i$.

Since $t_n \downarrow 0$, we see that T is compact. Furthermore, if $x \in K$, then let $y = (y_i)_{i=1}^{\infty} \in \ell^2$ with $y_i = \frac{x_i}{t_i}$, $i \in \mathbb{N}$. Then $Ty = x$, and

$$\sum_{i=1}^{\infty} y_i^2 = \sum_{i=1}^{\infty} \left(\frac{x_i}{t_i}\right)^2 = \sum_{i=1}^{N_1-1} \left(\frac{x_i}{t_i}\right)^2 + \sum_{n=1}^{\infty} \sum_{i=N_n}^{N_{n+1}-1} \left(\frac{x_i}{t_i}\right)^2$$

with

$$\sum_{i=N_n}^{N_{n+1}-1} \left(\frac{x_i}{t_i}\right)^2 = 2^{n-1} \sum_{i=N_n}^{N_{n+1}-1} x_i^2 \leq 2^{-n-1} \quad \text{and} \quad \sum_{i=1}^{N_1-1} \left(\frac{x_i}{t_i}\right)^2 \leq \frac{1}{2}.$$

We may conclude that

$$\sum_{i=1}^{\infty} y_i^2 \leq \frac{1}{2} + \sum_{n=1}^{\infty} 2^{-n-1} = 1,$$

so $y \in B(0,1)$. It follows that $K \subset T(B(0,1))$. □

Suppose Hypothesis 4.1 holds and consider, for $x \in H$, the stochastic variable

$$Y_x := \int_0^1 S(1-s)G(X(s,x))\, dM(s).$$

Lemma 4.7. *For all $\varepsilon > 0$ and $r > 0$, there exists a compact $K(r,\varepsilon)$ such that*

$$\mathbb{P}\left(Y_x \in K(r,\varepsilon)\right) > 1 - \varepsilon$$

for all $|x| \leq r$.

PROOF: Recall the factorization $G = C_2\Psi$ through the Banach space E_2 from Hypothesis 4.1 with C_2 compact. By Lemma 4.4, it is shown that if we let

$$V = \left\{S(t)C_2k : t \in [0,1], k \in E_2, |k| \leq 1\right\},$$

and K the closed convex hull of V, then K is compact. Let $T \in L(H)$, compact, be as given by Lemma 4.6, so $K \subset T(B(0,1))$, and let $K(\lambda) := \lambda T(B(0,1))$ for $\lambda > 0$, where $B(0,1)$ is the unit ball in H.

Note that

$$\begin{aligned} Y_x &= \int_0^1 S(1-s)C_2\Psi(X(s,x))\, dM(s) \\ &= \int_0^1 TT^{-1}S(1-s)C_2\Psi(X(s,x))\, dM(s) \\ &= T\int_0^1 T^{-1}S(1-s)C_2\Psi(X(s,x))\, dM(s). \end{aligned}$$

So

$$Y_x \in K(\lambda) \Leftrightarrow \int_0^1 T^{-1}S(1-s)C_2\Psi(X(s,x))\, dM(s) \in \lambda B(0,1).$$

Hence, using the fact that $T^{-1}S(1-s)C$ is an operator of norm not greater than 1 (by definition of T),

$$
\begin{aligned}
\mathbb{P}(Y_x \notin K(\lambda)) &\leq \mathbb{P}\left(\int_0^1 T^{-1}S(1-s)C_2\Psi(X(s,x))\ dM(s) \notin \lambda B(0,1)\right) \\
&\leq \frac{1}{\lambda^2}\mathbb{E}\left[\left|\int_0^1 T^{-1}S(1-s)C_2\Psi(X(s,x))\ dM(s)\right|^2\right] \\
&= \frac{1}{\lambda^2}\mathbb{E}\left[\int_0^1 \left|T^{-1}S(1-s)C_2\Psi(X(s,x))\right|^2_{L_{\mathrm{HS}}(\mathcal{H};E_2)}\ ds\right] \\
&\leq \frac{1}{\lambda^2}\mathbb{E}\left[\int_0^1 |\Psi(X(s,x))|^2_{L_{\mathrm{HS}}(\mathcal{H};E_2)}\ ds\right] \leq \frac{c_2}{\lambda^2}(1+|x|^2),
\end{aligned}
$$

for some constant $c > 0$, and where we used the second estimate of Theorem 2.27 in the last step. Now pick λ large enough such that

$$\frac{c_2}{\lambda^2}(1+r^2) < \varepsilon.$$

□

Lemma 4.8. *Suppose Hypothesis 4.1 is satisfied. For any $\varepsilon > 0$ and $r > 0$ there exists a compact $K(r,\varepsilon) \subset H$ such that*

$$\mathbb{P}(X(1,x) \in K(r,\varepsilon)) \geq 1-\varepsilon \quad \textit{for all } x \in H \textit{ with } |x| \leq r.$$

Proof: Note that

$$X(1,x) = S(1)x + \int_0^1 S(1-s)F(X(s,x))\ ds + \int_0^1 S(1-s)G(X(s,x))\ dM(s).$$

We treat the three terms separately.

Since $S(1)$ is a compact operator, for any $r > 0$ there exists a compact set $K_1(r)$ such that $S(1)x \in K_1(r)$ for all $|x| \leq r$.

Let $p \geq 2$. From the second estimate of Theorem 2.27, it follows that there exists a constant $k > 0$ such that

$$\mathbb{E}\left[\int_0^1 |\Phi(X(s,x))|^p\ ds\right] \leq k(1+|x|^p).$$

Define

$$f : \Omega \times [0,1] \to E_1, \quad f(t) := \Phi(X(t,x)), \quad t \in [0,1].$$

Then for $\lambda > 0$,

$$\mathbb{P}(|f|_{L^p([0,1];E_1)} > \lambda) \leq \frac{1}{\lambda^p}\mathbb{E}\left[|f|^p_{L^p([0,1];E_1)}\right] \leq \frac{k}{\lambda^p}(1+|x|^p) \leq \frac{k}{\lambda^p}(1+r^p).$$

Pick

$$\lambda := \left(\frac{2k}{\varepsilon}(1+r^p)\right)^{1/p},$$

so that

$$\mathbb{P}(|f|_{L^p([0,1];E_1)} > \lambda) \leq \varepsilon/2.$$

Hence, by Lemma 4.5 there exists a compact set $K_2(\lambda) = K_2(r,\varepsilon)$ such that

$$\mathbb{P}\left(\int_0^1 S(1-s)F(X(s,x))\ ds \in K_2(r,\varepsilon)\right) > 1-\varepsilon/2.$$

By Lemma 4.7, there exists a compact set $K_3(r,\varepsilon)$ such that

$$\mathbb{P}\left(\int_0^1 S(1-s)G(X(s,x))\ ds \in K_3(r,\varepsilon)\right) > 1-\varepsilon/2.$$

We may conclude that

$$\mathbb{P}(X(1,x) \in K_1(r) + K_2(r,\varepsilon) + K_3(r,\varepsilon)) \geq 1-\varepsilon.$$

□

Proof of Theorem 4.3: The proof is analogous to the proof of [25], Theorem 6.1.2.

Let $K(r,\varepsilon)$ as in Lemma 4.8. For $t > 1$, using Markov transition probabilities $(p_t(x,\ dy))$,

$$\begin{aligned}\mathbb{P}(X(t,x) \in K(r,\varepsilon)) &= \mathbb{E}\left[p_1(X(t-1,x), K(r,\varepsilon))\right] \\ &\geq \mathbb{E}\left[p_1(X(t-1,x), K(r,\varepsilon))\mathbb{1}_{\{|X(t-1,x)|\leq r\}}\right].\end{aligned}$$

By Lemma 4.8,

$$\mathbb{P}(X(t,x) \in K(r,\varepsilon)) \geq (1-\varepsilon)\mathbb{P}(|X(t-1,x)| \leq r),$$

so

$$\frac{1}{T}\int_1^{T+1} \mathbb{P}(X(t,x) \in K(r,\varepsilon))\ dt \geq \frac{1-\varepsilon}{T}\int_0^T \mathbb{P}(|X(t,x)| \leq r)\ dt.$$

Now, using Hypothesis 4.1, (vii), take r large enough and ε small enough, to see that the family

$$\frac{1}{T}\int_1^{T+1} p_t(x,\cdot)\ dt, \quad T \geq 1,$$

is tight. By Krylov-Bogoliubov there exists an invariant measure. □

Remark 4.9. We conjecture that a sufficient condition for existence of invariant measures is compactness of $T(t)C_1$ and $T(t)C_2$ for all $t > 0$. By letting $C = I$ this would include the case proven in [22] for compact semigroups S and without conditions on F and G. Compactness of the drift term in the case where $T(t)C_1$ is compact for $t > 0$ is shown in Lemma 4.5. However a way of proving tightness of the noise term for $T(t)C_2$ compact for all $t > 0$ escapes us so far.

4.1.1 Boundedness in probability

We state here a sufficient condition for boundedness in probability (based on [22], Proposition 7). The proof is identical to that of the cited proposition, except that here the semigroup is not assumed to be immediately Hilbert-Schmidt, but instead G maps into the space of Hilbert-Schmidt mappings. This way the result below is applicable to the case of stochastic delay differential equations, where the semigroup is not immediately compact.

Proposition 4.10. *Let M be a cylindrical martingale with RKHS $\mathcal{H}$. Suppose A generates a strongly continuous semigroup $(T(t))_{t\geq 0}$ such that*

$$\int_0^\infty ||T(t))||^2 \; dt < \infty.$$

Moreover let $F : H \to H$ satisfy

$$\langle Ax + F(x+y), x\rangle \leq a + a|y|^2 - 2b|x|^2, \quad x \in \mathfrak{D}(A),$$

with a and b positive constants, and let $G : H \to L_{\mathrm{HS}}(\mathcal{H}; H)$ be bounded, i.e. $\sup_{x\in H} ||G(x)||_{L_{\mathrm{HS}}(\mathcal{H};H)} < \infty$.

Then

$$\sup_{t\geq 0} \mathbb{E}|X(t;x)|^2 < \infty,$$

where $(X(t;x))_{t\geq 0}$ is the unique mild solution of (4.1).

4.2 Application to stochastic evolutions with delay

Let $(S(t))_{t\geq 0}$ be an abstract delay semigroup as introduced in Section 3.1.

By Theorem 3.16 we see that in many cases A generates an eventually compact semigroup. We apply this result, in combination with the existence result of the previous section, to establish the existence of an invariant measure for two particular types of stochastic evolutions with delay.

4.2.1 Example: Functional differential equations

A relatively easy case is now the example of a functional differential equation perturbed by noise given in Section 3.3,

$$dx(t) = \left[Bx(t) + \sum_{i=1}^k B_i x(t - h_i) + \varphi(x(t), x_t) \right] \; dt + \psi(x(t), x_t) \; dM(t). \tag{4.2}$$

Since $F = \begin{pmatrix} \varphi \\ 0 \end{pmatrix}$ and $G = \begin{pmatrix} \psi \\ 0 \end{pmatrix}$ map into finite dimensional subspaces of $\mathcal{E}^2$ and $L(\mathbb{R}^k; \mathcal{E}^2)$, they admit a compact factorization as demanded in Hypothesis 4.1. By Theorem 3.16, A generates an eventually compact semigroup. Hence all conditions of Hypothesis 4.1 are satisfied, apart from condition (vi), boundedness in probability, which remains to be verified. If this condition is also satisfied, by Theorem 4.3 we have established the existence of an invariant measure for (4.2) on the state space $\mathcal{E}^2$.

4.2.2 Example: Reaction-diffusion equations with noise and delay

Reaction-diffusion equations with delayed nonlocal reaction terms are a topic of active research in the study of biological invasion and disease spread. Can we establish the existence of an invariant measure if we add randomness to such a system? As an example we set out to answer this question for an equation similar to one encountered in e.g. [53].

Consider the reaction-diffusion equation with delay and noise

$$\begin{cases} du(t,\xi) & \\ = \left[\Delta_\xi u(t,\xi) + \sum_{i=1}^n c_i \frac{\partial}{\partial \xi} u(t-\theta_i,\xi) + \varphi(u_t)(\xi)\right] \, dt & t \geq 0, \xi \in \mathbb{R}, \\ \quad +(\sigma \circ dM)(t,\xi), & \\ \lim_{\xi \to \pm\infty} u(t,\xi) = \lim_{\xi \to \pm\infty} \frac{\partial u}{\partial \xi}(t,\xi) = 0, & t \geq 0, \\ u(t,\xi) = f(t,\xi), & t \in [-1,0], \xi \in \mathbb{R}, \\ u(0,\xi) = v(\xi), & \xi \in \mathbb{R}. \end{cases} \tag{4.3}$$

with

(i) delay parameters $c_i \in \mathbb{R}, \theta_i \in [-1,0]$, $i = 1, \ldots, n$,

(ii) initial conditions $f \in L^2([-1,0]; W^{1,2}(\mathbb{R}))$ and $v \in L^2(\mathbb{R})$,

(iii) Lipschitz reaction term $\varphi : L^2([-1,0]; W^{1,2}(\mathbb{R})) \to L^2(\mathbb{R})$ (possibly non-linear and/or non-local),

(iv) $u_t \in L^2([-1,0]; W^{1,2}(\mathbb{R}))$ denoting the segment process defined by $u_t(\theta) = u(t+\theta)$ for $\theta \in [-1,0]$, $t \geq 0$,

(v) $(M(t))_{t \geq 0}$ a cylindrical martingale with RKHS $\mathcal{H}$,

(vi) noise factor $\sigma \in L_{\mathrm{HS}}(\mathcal{H}; W^{1,2}(\mathbb{R}))$.

We can employ the semigroup approach discussed before by setting $X = L^2(\mathbb{R})$, $Z := W^{1,2}(\mathbb{R})$, as state space the Hilbert space $\mathcal{E}^2 = L^2(\mathbb{R}) \times L^2([-1,0]; W^{1,2}(\mathbb{R}))$, with A as defined in (3.2), with

$$B := \Delta, \quad \mathfrak{D}(B) = \left\{ v \in W^{2,2}(\mathbb{R}) : \lim_{\xi \to \pm\infty} v(\xi) = 0, \lim_{\xi \to \pm\infty} \frac{\partial v}{\partial \xi}(\xi) = 0 \right\}$$

and

$$\Phi(w) := \sum_{i=1}^{n} c_i \frac{\partial}{\partial \xi} w(t - \theta_i, \xi), \quad w \in L^2([-1,0]; W^{1,2}(\mathbb{R})).$$

Then A is of the form described in Example 3.6. Since B generates an immediately compact semigroup, it follows from Theorem 3.16 that A generates an eventually compact semigroup.

Furthermore let

$$F\left(\begin{bmatrix} v \\ w \end{bmatrix}\right) := \begin{bmatrix} \varphi(w) \\ 0 \end{bmatrix}, \quad \begin{bmatrix} v \\ w \end{bmatrix} \in \mathcal{E}^2, \quad \text{and} \quad G(\cdot) := \begin{bmatrix} \sigma \\ 0 \end{bmatrix} \quad \text{on } \mathcal{E}^2.$$

Then (4.3) is described by (4.1) in the state space $H = \mathcal{E}^2$.

It remains to impose conditions on the nonlinear term F. Let us require, for example, that $\varphi : L^2([-1,0]; W^{1,2}(\mathbb{R})) \to L^2(\mathbb{R})$ is of the form

$$\varphi(w) := (g \circ h)(w),$$

with $g : W^{1,2}(\mathbb{R}) \to W^{1,2}(\mathbb{R})$ (possibly the identity mapping), and h defined by

$$h(w)(\xi) := \int_{-1}^{0} \int_{\mathbb{R}} k(\eta, \sigma) \psi(w(\sigma, \xi - \eta)) \, d\eta \, d\sigma, \quad \xi \in \mathbb{R},$$

where $\psi \in W^{1,\infty}(\mathbb{R})$ with $|\psi(\zeta)| \leq ||\dot{\psi}||_\infty |\zeta|$, $\zeta \in \mathbb{R}$, and $k \in L^1(\mathbb{R}; L^2([-1,0]))$.

We will now verify that in this case

$$\varphi : L^2([-1,0]; W^{1,2}(\mathbb{R})) \to W^{1,2}(\mathbb{R}). \tag{4.4}$$

Indeed, using Fubini, Cauchy-Schwarz and Young's inequality for convolutions,

$$\begin{aligned}
\int_{\mathbb{R}} |h(w)(\xi)|^2 \, d\xi &= \int_{\mathbb{R}} \left| \int_{-1}^{0} \int_{\mathbb{R}} k(\eta, \sigma) \psi(w(\sigma, \xi - \eta)) \, d\eta \, d\sigma \right|^2 \, d\xi \\
&\leq ||\dot{\psi}||_\infty^2 \int_{\mathbb{R}} \left(\int_{-1}^{0} \int_{\mathbb{R}} |k(\eta, \sigma) w(\sigma, \xi - \eta)| \, d\eta \, d\sigma \right)^2 \, d\xi \\
&\leq ||\dot{\psi}||_\infty^2 \int_{\mathbb{R}} \left(\int_{\mathbb{R}} ||k(\eta, \cdot)||_{L^2([-1,0])} ||w(\cdot, \xi - \eta)||_{L^2([-1,0])} \, d\eta \right)^2 \, d\xi \\
&\leq \left(||\dot{\psi}||_\infty ||k||_{L^1(\mathbb{R}; L^2([-1,0]))} \; ||w||_{L^2([-1,0]; L^2(\mathbb{R}))} \right)^2.
\end{aligned}$$

Furthermore, using the same classic inequalities,

$$
\begin{aligned}
&\int_{\mathbb{R}} \left| \frac{\partial}{\partial \xi} h(w)(\xi) \right|^2 \, d\xi \\
&= \int_{\mathbb{R}} \left| \int_{-1}^{0} \int_{\mathbb{R}} k(\eta,\sigma) \frac{\partial}{\partial \xi} \psi(w(\sigma, \xi - \eta)) \, d\eta \, d\sigma \right|^2 \, d\xi \\
&= \int_{\mathbb{R}} \left| \int_{-1}^{0} \int_{\mathbb{R}} k(\eta,\sigma) \dot{\psi}(w(\sigma, \xi - \eta)) \frac{\partial}{\partial \xi} w(\sigma, \xi - \eta) \, d\eta \, d\sigma \right|^2 \, d\xi \\
&\leq ||\dot{\psi}||_\infty^2 \int_{\mathbb{R}} \left(\int_{-1}^{0} \int_{\mathbb{R}} \left| k(\eta,\sigma) \frac{\partial}{\partial \xi} w(\sigma, \xi - \eta) \right| \, d\eta \, d\sigma \right)^2 \, d\xi \\
&\leq ||\dot{\psi}||_\infty^2 \int_{\mathbb{R}} \left(\int_{\mathbb{R}} ||k(\eta, \cdot)||_{L^2([-1,0])} \left|\left| \frac{\partial}{\partial \xi} w(\cdot, \xi - \eta) \right|\right|_{L^2([-1,0])} \, d\eta \right)^2 \, d\xi \\
&\leq \left(||\dot{\psi}||_\infty ||k||_{L^1(\mathbb{R}; L^2([-1,0]))} \, ||w||_{L^2([-1,0]; W^{1,2}(\mathbb{R}))} \right)^2 .
\end{aligned}
$$

So we have $h : L^2([-1,0]; L^2(\mathbb{R})) \to W^{1,2}(\mathbb{R})$, and therefore (4.4) holds for $\varphi = g \circ h$. Hence in this case we may write (with some abuse of notation) $\varphi = \imath \circ \varphi$, where $\imath : W^{1,2}(\mathbb{R}) \to L^2(\mathbb{R})$ is the canonical embedding of $W^{1,2}(\mathbb{R})$ into $L^2(\mathbb{R})$, which is a compact mapping. We conclude that F admits a compact factorization. Note that this carries over to any function φ that satisfies (4.4). G admits a compact factorization as well, again using the compact embedding of $W^{1,2}(\mathbb{R})$ into $L^2(\mathbb{R})$.

Again, we may conclude from Theorem 4.3 that if the solutions of (4.3) are bounded in probability, an invariant measure exists.

4.2.3 Example: Environmental pollution

In the example of Section 3.4.2,

$$
\begin{cases}
du(t,\xi) & = (\gamma \Delta u(t,\xi) - \kappa u(t,\xi) + \eta u(t-1,\xi) \, dt \\
& \qquad + (\rho \star dZ(t))(\xi) \qquad\qquad t > 0, \xi \in \mathcal{O}, \\
\frac{\partial u(t,\xi)}{\partial \nu} & = 0, \qquad\qquad\qquad\qquad t > 0, \xi \in \partial\mathcal{O}, \\
u(0,\xi) & = x(\xi), \qquad\qquad\qquad\quad \xi \in \mathcal{O}, \\
u(t,\xi) & = f(t,\xi), \qquad\qquad\qquad -1 \leq t \leq 0, \xi \in \mathcal{O}.
\end{cases}
\tag{4.5}
$$

with as special choice $\rho \in W^{1,2}(\mathcal{O})$, it may easily be verified that the convolution $\rho \star u$ maps into $W^{1,2}(\mathcal{O})$. We may again use the compact embedding of $W^{1,2}(\mathcal{O}) \hookrightarrow L^2(\mathcal{O})$ to obtain a compact factorization of the noise term.

So if the solution of (4.5) is bounded in probability, we may conclude that an invariant measure exists. We may also perturb the stochastic differential equation by a nonlinear term, provided that condition (iii) of Hypothesis 4.1 (which requires the perturbation to admit a compact factorization) is satisfied.

4.3 Dissipative systems

We briefly mention here a result in [71], which is an extension of Theorem 6.3.2 of [25] to the case of Lévy processes. See Section 6.4.2 for a short introduction to Yosida approximations.

Theorem 4.11. *Suppose A is the generator of a strongly continuous semigroup $(T(t))_{t\geq 0}$ in H , $F : H \to H$ is globally Lipschitz, L a cylindrical mean-zero Lévy process with RKHS $\mathcal{H}$ and $G : H \to L_{\mathrm{HS}}(\mathcal{H}; H)$ is globally Lipschitz. Furthermore suppose there exists $\omega > 0$ such that*

$$2\langle A_n(x-y) + F(x) - F(y), x-y\rangle + ||G(x) - G(y)||^2_{L_{\mathrm{HS}}(\mathcal{H};H)} \leq -\omega|x-y|^2$$

for all $x, y \in H$ and $n \in \mathbb{N}$, where $(A_n)_{n\in\mathbb{N}}$ are the Yosida approximations of A.

Then there exists exactly one invariant measure μ for the stochastic differential equation

$$\begin{cases} dX(t) = [AX(t) + F(X(t)]\ dt + G(X(t))\ dL(t), & t \geq 0 \\ X(0) = x \in H, \end{cases}$$

it is strongly mixing and there exists $C > 0$ such that for any bounded Lipschitz continuous function φ, all $t > 0$ and $x \in H$,

$$\left| P(t)\varphi(x) - \int_H \varphi\ d\mu \right| \leq C(1 + |x|)\mathrm{e}^{-\omega t/2}||\varphi||_{\mathrm{Lip}}.$$

PROOF: See [71], Theorem 16.5. □

This result is particularly interesting in light of Section 3.6, in which we show how the infinitesemal generator of the delay semigroup can be made into a dissipative operator.

4.3.1 Example: Environmental pollution

Recall the environmental pollution model (3.11) introduced in Section 3.4.2.

Proposition 4.12. *If $0 \leq \eta < \kappa$, then there exists a unique invariant measure which is strongly mixing for the evolution described by* (3.11)*.*

PROOF: In this case, we may find a small $\omega > 0$, $\omega < \kappa$, such that

$$(\kappa - \omega) > \eta \mathrm{e}^{\omega}.$$

We may therefore apply Theorem 3.18 with $\lambda = -\kappa$ and $\mu = -\omega$ to obtain that for A defined in (3.12) we have that $\langle Ax, x\rangle \leq -\omega\langle x, x\rangle$ (under a new equivalent inner product).

By (6.16) of Section 6.4.2, the same holds for the Yosida approximants, for n large enough.

We may now apply Theorem 4.11 to find that the evolution described by (3.11) admits a unique invariant measure which is strongly mixing. $\square$

Note that the intensity of the noise has no effect due to its additive character.

4.4 Notes and remarks

Parts of this chapter have been accepted for publication as [12].

The literature on existence (and uniqueness) of invariant measures of stochastic evolutions is vast. To mention a few, results on existence and uniqueness of invariant measures are given in [19], [18], [92]. For delay differential equations with Lévy noise see [35], [89] and [74].

In [84] existence of invariant measures for affine stochastic differential equations in some Banach spaces with a semigroup which is not asymptotically stable is discussed.

CHAPTER

FIVE

Uniqueness of invariant probability measure for degenerate evolutions

A stochastic delay differential equation with additive noise can be modeled (see Section 3.2) as a stochastic Cauchy problem in some Hilbert space H of the form

$$\begin{cases} dX(t) = [AX + F(X)]\ dt + G\ dW(t) & t \geq 0, \text{a.s.} \\ X(0) = x \quad \text{a.s.} \end{cases} \tag{5.1}$$

where A is the generator of the delay semigroup, F a sufficiently smooth function (e.g. Lipschitz), and G a linear operator mapping the Wiener process W into H. It is well known that under the mentioned assumptions, existence and uniqueness of solutions is guaranteed.

So far however, the ergodic behaviour of these systems was less well understood. An important notion in this respect is that of *invariant probability measure*, i.e. a positive finite Borel measure μ on H with $\mu(H) = 1$ such that if the initial condition x has law μ, then the solution $X(t; x)$ has law μ for all $t \geq 0$. Recently the existence of an invariant (probability) measure was established for a sufficiently broad class of stochastic Cauchy problems to include the case of finite dimensional stochastic delay differential equations [12]; see also Chapter 4.

Apart from the existence of an invariant probability measure, its uniqueness is an important issue. When an invariant probability measure is unique, the *ergodic property* 'time average equals spatial average' holds:

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T \varphi(X(t; x))\ dt = \int_H \varphi\ d\mu, \quad \varphi \in B_b(H),$$

where $B_b(H)$ are the bounded Borel measurable functions on H.

Just as the problem of existence of invariant measures, also the problem of uniqueness of the invariant probability measure of stochastic delay differential equations were open for some time. A partial solution to this problem was proposed by using the dissipativity properties of the delay semigroup (see [25] and sections 3.6 and 4.3).

In [70] general conditions for the uniqueness of the invariant probability measure are established for the nondegenerate noise case. However, the noise that perturbs delay equations can influence only the present of the process and not the past and is therefore essentially degenerate, so these results do not apply here. In [19] results are obtained for degenerate noise, but these do not include the case of delay equations.

Often uniqueness of invariant probability measure is proved using Doob's theorem (see e.g. [25], Theorem 5.2.1). This requires irreducibility and the strong Feller property of solutions. In [25] the eventual strong Feller property for systems of the form (5.1) was conjectured. This property states that $P(t)\varphi$ is continuous and bounded for any $\varphi \in B_b$ and is important in establishing the uniqueness of the invariant probability measure. It is not immediate that the strong Feller property holds, because usually some kind of non-degeneracy assumption on the noise is required. However, in the case of stochastic delay differential equations, the noise is intrinsically degenerate because it can only work on the 'present' of the process, while the state space also contains the 'past' of the stochastic evolution.

In this chapter we establish conditions that are sufficient to establish uniqueness of the invariant probability measure for degenerate stochastic Cauchy problems of the form (5.1). We combine methods from the now classical semigroup approach initiated by Da Prato and Zabczyk [24], and from Malliavin calculus, inspired by succesful applications in e.g. [37], to obtain the eventualy strong Feller property and eventual irreducibility. In [59] the eventual strong Feller property for delay equations with additive noise was also established by a probabilistic method. However, we think the conditions established by our method are easier to verify in practice.

Our main result is stated in Section 5.1, and the proof is split into two parts, discussed in Sections 5.2 (strong Feller property) and 5.3 (irreducibility). The result is applied to stochastic delay differential equations and a stochastic partial differential equation with delay in Section 5.4.1.

Recently, Hairer and Mattingly [37] introduced a generalization of the strong Feller property which is still strong enough to help in establishing uniqueness of invariant measure. We study this property in relation to controllability and stability properties of the linear part of the equation in Section 5.5.

5.1 Main result

We need the notions of null controllability, approximate controllability and the Fréchet derivative.

Let $\mathcal{H}, H$ be Hilbert spaces. Consider the controlled Cauchy problem

$$\begin{cases} \dot{x}(t) = Ax(t) + f(x(t)) + Gu(t), & t \geq 0 \\ x(0) = x_0 \end{cases} \tag{5.2}$$

with A the generator of a strongly continuous semigroup $(S(t))_{t\geq 0}$ on H, $f : H \to H$ globally Lipschitz, $G \in L(\mathcal{H}; H)$ and where $u \in L^2([0,T]; \mathcal{H})$ is called the *control*.

Definition 5.1 (null controllability)**.** *The system (5.2) is* null controllable *in time $t > 0$ if for any $x_0 \in H$ there exists a control $u \in L^2([0,t]; \mathcal{H})$ such that $x(t; x_0) = 0$.*

The pair (A, G) is called null controllable *in time $t > 0$ if (5.2) with $f \equiv 0$ is null controllable in time $t > 0$.*

It is well known (see [24], Section B.3) that null controllability of (A, G) in time $t > 0$ is equivalent to

$$\text{im } S(t) \subset \text{im } Q_t^{1/2},$$

where the *controllability Grammian* $Q_t \in L(H)$ is defined by

$$Q_t x := \int_0^t S(s)GG^*S^*(s)x \, ds, \quad x \in H. \tag{5.3}$$

Furthermore, since the linear operator $Q_t^{-1/2}S(t) : H \to H$ is closed and defined everywhere on H, by the closed graph theorem it is bounded.

Definition 5.2 (approximate controllability)**.** *The system (5.2) is said to be* approximately controllable *in time $t > 0$ if, for arbitrary $x_0, z \in H$ and $\varepsilon > 0$, there exists a control $u \in L^2([0,t]; \mathcal{H})$ such that $|x(t; x_0, u) - z| < \varepsilon$.*

The pair (A, G) is said to be approximately controllable *in time $t > 0$ if (5.2) with $f \equiv 0$ is approximately controllable in time $t > 0$.*

Suppose H, K are Hilbert spaces and $F : H \to K$ is Fréchet differentiable. We then denote the Fréchet differential of F by $dF : H \to L(H; K)$. Let V denote a closed subspace of K containing im (F) and note that $dF : H \to L(H; V)$.

The following Hypothesis states conditions sufficient to prove the uniqueness of the invariant probability measure.

Hypothesis 5.3. *(i) H is a Hilbert space;*

(ii) $(S(t))_{t\geq 0}$ is a strongly continuous semigroup acting on H with generator A;

(iii) W is a cylindrical Wiener process with RKHS $\mathcal{H}$;

(iv) $F : H \to V \subset \text{im } (G)$, with V a closed subspace of H, is globally Lipschitz;

(v) $G \in L(\mathcal{H}; H)$ and a mapping $G^{-1} \in L(V, \mathcal{H})$ exists such that $GG^{-1} = I$ on V;

(vi) *The pair (A, G) is both null controllable and approximately controllable at $T > 0$, and $Q := Q_T$ is defined by (5.3).*

In many cases it is convenient to take $V = \text{im } (G)$. However if F maps into a strict subspace of im (G) the condition of pseudoinvertibility of G can be relaxed by letting $V \subsetneq \text{im } (G)$.

We will assume throughout this section that for any $x \in H$, there exists a unique mild solution $(X(t;x))_{t \geq 0}$ of (5.1). Sufficient conditions for this to hold are that $G \in L_{\text{HS}}(\mathcal{H}; H)$ (see Theorem 2.27).

Before we can state the main result of this chapter, we introduce some more notions. The *(Markov) transition semigroup* associated to a Markov process $(X(t;x))$ is defined as the family of operators $(P(t))_{t \geq 0}$ acting on $B_b(H)$, defined by

$$P(t)\varphi(x) = \mathbb{E}[\varphi(X(t;x))], \quad \varphi \in B_b(H), x \in H, t \geq 0.$$

The transition semigroup $(P(t))_{t \geq 0}$ is called *strong Feller* at $t > 0$ if $P(t)\varphi \in C_b(H)$ for all $\varphi \in B_b(H)$, and *irreducible* at $t > 0$ if $P(t)\mathbb{1}_\Gamma(x) > 0$ for any open, non-empty $\Gamma \subset H$, $x \in H$. A positive Borel measure μ on H is said to be *invariant* for $(P(t))_{t \geq 0}$ if

$$\int_H P(t)\varphi \, d\mu = \int_H \varphi \, d\mu \quad \text{for all } \varphi \in B_b(H), t \geq 0.$$

If furthermore $\mu(H) = 1$ then μ is called *invariant probability measure*.

An invariant measure μ is called *strongly mixing* if

$$\lim_{t \to \infty} P(t)\varphi(x) = \int_H \varphi \, d\mu, \quad \text{for all } \varphi \in B_b(H), x \in H.$$

The following theorem is the main result of this chapter.

Theorem 5.4. *Suppose the assumptions of Hypothesis 5.3 hold. Then there exists at most one invariant probability measure for (5.1), and if it exists, this invariant measure is strongly mixing.*

PROOF: By Theorem 5.10, the transition semigroup associated to (5.1) is strong Feller at time T, and by Corollary 5.13 it is irreducible at time T. Hence by Khas'minskii's theorem ([25], Theorem 4.1.1), the transition semigroup of (5.1) is regular at time $2T$. Then the conclusion follows from Doob's theorem ([25], Theorem 4.2.1). □

5.2 Null controllability and the strong Feller property

The following hypothesis is almost similar to Hypothesis 5.3. Only condition (vi) is slightly weaker: we do not need approximate controllability. We will see that the hypothesis is sufficient to prove the strong Feller property of (5.1).

Hypothesis 5.5. *(i) H is a Hilbert space;*

(ii) $(S(t))_{t\geq 0}$ is a strongly continuous semigroup acting on H with generator A;

(iii) W is a cylindrical Wiener process with RKHS $\mathcal{H}$;

(iv) $F : H \to V \subset \text{im }(G)$, with V a closed subspace of H, is globally Lipschitz;

(v) $G \in L(\mathcal{H}; H)$ and a mapping $G^{-1} \in L(V, \mathcal{H})$ exists such that $GG^{-1} = I$ on V;

(vi) The pair (A, G) is null controllable at $T > 0$, and $Q := Q_T$ is defined by (5.3);

5.2.1 Linearized flow

We are interested in dependence of the solution $(X(t; x))_{t\geq 0}$ of (5.1) on the initial condition x. Therefore we define, for arbitrary directions $\xi \in H$, the derivative processes $J_{0,t}\xi := d_x X(t; x)\xi$, where d_x is the Fréchet differential of $X(t; x)$ with respect to x. Assume for now that $F : H \to \text{im }(G)$ is continuously Fréchet differentiable, with $||dF||_\infty < \infty$.

By [25], Theorem 5.4.1, $J_{0,t}\xi$ is a mild solution to

$$\begin{cases} \frac{d}{dt} J_{0,t}\xi = AJ_{0,t}\xi + dF(X(t; x))J_{0,t}\xi & \text{a.s., } t \geq 0 \\ J_{0,s} = \xi \quad \text{a.s.,} & \end{cases}$$

and there exists a constant $C > 0$ independent of ξ such that

$$\sup_{t\in[0,T]} \mathbb{E}|J_{0,t}\xi|^2 \leq C|\xi|^2. \tag{5.4}$$

More generally define for $s \geq 0$ and $t \geq s$ the linear, stochastic operators $J_{s,t}$ as the pathwise solutions of

$$J_{s,t}\xi = S(t - s)\xi + \int_s^t S(t - r)dF(X(r; x))J_{s,r}\xi \, dr \tag{5.5}$$

for $\xi \in H$.

We set out to express the dependence of $X(T; x)$ on the initial condition x in terms the dependence of $X(T; x)$ on the noise process W. For this we need the notion of Malliavin derivative.

5.2.2 Malliavin calculus

Our exposition of the Malliavin calculus is based on [16], Chapter 5.

Let W be a cylindrical Brownian motion with reproducing kernel Hilbert space $\mathcal{H}$ and let K be a separable Hilbert space.

We first define the Malliavian derivative of smooth variables. A random variable $X \in L^2(\Omega; K)$ is called *smooth* if X has the form

$$X = \psi(W(\Phi_1), \ldots, W(\Phi_n)),$$

with $\psi : \mathbb{R}^n \to K$ infinitely often differentiable, $\Phi_1, \ldots, \Phi_n \in L^2([0,T]; \mathcal{H})$ and

$$W(\Phi) := \int_0^T \langle \Phi(t),\ dW(t) \rangle, \quad \Phi \in L^2([0,T]; \mathcal{H}).$$

We denote all smooth K-valued random variables by $\mathcal{S}(K)$.

For $X \in \mathcal{S}(K)$ we define the *Malliavin derivative* DX of X as the $K \otimes L^2([0,T]; \mathcal{H})$-valued random variable

$$DX = \sum_{i=1}^n \frac{\partial}{\partial x_i} \psi(W(\Phi_1), \ldots, W(\Phi_n)) \otimes \Phi_i.$$

Note that we may identify the range of D with $L^2(\Omega \times [0,T]; L_{\mathrm{HS}}(\mathcal{H}; K))$, so we can (and will) interpret DX as a (possibly non-adapted) stochastic process $(D_t X)_{t \in [0,T]}$ with values in $L_{\mathrm{HS}}(\mathcal{H}; K)$.

The mapping $X \mapsto DX : \mathcal{S}(K) \to L^2(\Omega \times [0,T]; L_{\mathrm{HS}}(\mathcal{H}; K))$ is closable ([16], Proposition 5.1), and we call its closure $D : \mathbb{H}(K) \to L^2(\Omega \times [0,T]; L_{\mathrm{HS}}(\mathcal{H}; K))$ the *Malliavin derivative*, where the domain $\mathbb{H}(K)$ of D is a linear subspace of $L^2(\Omega; K)$.

For $v \in L^2(\Omega \times [0,T]; \mathcal{H})$ we define the *Malliavin derivative in the direction* v pointwise almost everywhere on Ω as the K-valued square integrable random variable

$$D^v X := \int_0^T D_t X \circ v(t)\ dt.$$

Remark 5.6. Intuitively, DX is the stochastic process which, when integrated with respect to W over $[0,T]$, results in the random variable X. As such, DX represents the dependence of X on the noise process W, and $D^v X$ indicates the infinitesemal change in X if we perturb W infinitesemally in the direction of v. Note that this interpretation makes sense only if $(D_t X)_{t \in [0,T]}$ is adapted; see however the Skorohod integral below. ◇

We will use the following version of the *chain rule for the Malliavin derivative* (which holds more generally, see [16], Proposition 5.2): Suppose K_1 and K_2 are separable Hilbert spaces and assume $\varphi : K_1 \to K_2$ is Fréchet differentiable with uniformly bounded Fréchet derivative $d\varphi$. Then for $X \in \mathbb{H}(K_1)$, we have $\varphi(X) \in \mathbb{H}(K_2)$ and

$$D\varphi(X) = (d\varphi(X))(DX). \tag{5.6}$$

The adjoint operator $\delta : \mathfrak{D}(\delta, K) \to L^2(\Omega; K)$ of D is defined by the duality

$$\mathbb{E}\langle DX, \Phi\rangle_{L^2([0,T];L_{\mathrm{HS}}(\mathcal{H};K))} = \mathbb{E}\langle X, \delta\Phi\rangle_K,$$

for $X \in \mathbb{H}(K)$ and $\Phi \in \mathfrak{D}(\delta, K) \subset L^2(\Omega \times [0,T]; L_{\mathrm{HS}}(\mathcal{H}; K))$ and is called the *Skorohod integral*, also denoted by

$$\delta\Phi = \int_0^T \Phi(t)\ \delta W(t).$$

If Φ is a predictable process in $L_{\mathrm{HS}}(\mathcal{H}; K)$ such that

$$\mathbb{E}\int_0^T ||\Phi(t)||^2_{L_{\mathrm{HS}}(\mathcal{H};K)}\ dt < \infty,$$

then $\Phi \in \mathfrak{D}(\delta, K)$ and the Skorohod integral and the Itô integral coincide ([16], Theorem 5.1):

$$\int_0^T \Phi(t)\ \delta W(t) = \int_0^T \Phi(t)\ dW(t).$$

We therefore have, for predictable Φ, the integration by parts formula

$$\mathbb{E}\langle DX, \Phi\rangle_{K\otimes L^2([0,T];\mathcal{H})} = \mathbb{E}\left\langle X, \int_0^T \Phi(t)\ dW(t)\right\rangle_K$$

and in particular

$$\mathbb{E}\left[D^v X\right] = \mathbb{E}\left[X \int_0^T v(s)\ dW(s)\right]$$

where $X \in \mathbb{H}(K)$ and $v \in L^2(\Omega \times [0,T]; \mathcal{H})$ predictable.

We conclude our summary of Malliavin calculus with a commutation rule for the Malliavin derivative and the Skorohod integral (a straightforward extension to the infinite-dimensional case of [67], Proposition 1.3.2):

$$D^v\delta\Phi = \int_0^T \langle v(t), \Phi(t)\rangle_K\ dt + \delta(D^v\Phi), \tag{5.7}$$

which holds for (deterministic) $v \in L^2([0,T]; \mathcal{H})$ and $\Phi \in \mathfrak{D}(\delta, K)$ such that $D^v\Phi \in \mathfrak{D}(\delta)$.

Lemma 5.7. *Suppose* $(X(t;x))_{t\geq 0}$ *is the solution of*

$$\begin{cases} dX(t) = [AX(t) + F(X(t))]\ dt + G\ dW(t) & t \in [0,T] \\ X(0) = x, \end{cases}$$

with A *the generator of a strongly continuous semigroup* $(S(t))_{t\geq 0}$*,* $F : H \to H$ *Fréchet differentiable, and* $G \in L(\mathcal{H}; H)$*.*

Then, for $v \in L^2(\Omega \times [0,T]; \mathcal{H})$,

$$D^v X(t;x) = \int_0^t J_{s,t} G v(s)\ ds, \quad \text{almost surely}, t \in [0,T]$$

where $J_{s,t}$ is defined by (5.5).

PROOF: For $t > 0$ we have that $X(t;x) \in \mathbb{H}(H)$ by [16], Lemma 5.3. We have

$$D^v X(t;x) = D^v S(t)x + D^v \int_0^t S(t-s)F(X(s;x))\ ds + D^v \int_0^t S(t-s)G\ dW(s).$$

The first term disappears since $S(t)x$ is deterministic. By the chain rule of Malliavin calculus,

$$D \int_0^t S(t-s)F(X(s;x))\ ds = \int_0^t S(t-s)dF(X(s;x))DX(s;x)\ ds,$$

and hence

$$\begin{aligned}
&D^v \int_0^t S(t-s)F(X(s;x))\ ds \\
&= \int_0^T \int_0^t S(t-s)dF(X(s;x))D_r X(s;x)\ ds \circ v(r)\ dr \\
&= \int_0^T \int_0^t S(t-s)dF(X(s;x))D_r X(s;x) \circ v(r)\ dsdr \\
&= \int_0^t S(t-s)dF(X(s;x))D^v X(s;x)\ ds.
\end{aligned}$$

Finally by (5.7) for v deterministic

$$\begin{aligned}
D^v \int_0^t S(t-s)G\ dW(s) &= D^v \delta(S(t-s)G\mathbb{1}_{s\le t})_{s\in[0,T]} \\
&= \int_0^T S(t-s)G\mathbb{1}_{s\le t} v(s)\ ds = \int_0^t S(t-s)Gv(s)\ ds.
\end{aligned}$$

Hence for simple functions $v = \sum_{i=1}^n v_i \mathbb{1}_{E_i}$, with $E_i \in \mathcal{F}$ and $v_i \in L^2([0,T];\mathcal{H})$,

$$D^v \int_0^t S(t-s)G\ dW(s) = \int_0^t S(t-s)Gv(s)\ ds, \quad \text{almost surely.} \tag{5.8}$$

We obtain (5.8) for general $v \in L^2(\Omega \times [0,T]; \mathcal{H})$ by approximating v by simple functions.

Hence

$$D^v X(t;x) = \int_0^t S(t-s)dF(X(s;x))D^v X(s;x)\ ds + \int_0^t S(t-s)Gv(s)\ ds \quad \text{a.s.},$$

or equivalently, using the definition of $(J_{s,t})_{t \geq s}$ in (5.5),

$$D^v X(t;x) = \int_0^t J_{s,t} G v(s)\ ds, \quad t \in [0,T].$$

$\square$

Lemma 5.8. *Assume Hypothesis 5.5 holds and that* $F : H \to \mathrm{im}\ (G)$ *is Fréchet differentiable, with* $||dF||_\infty < \infty$.

Then for all $\xi \in H$ *there exists a stochastic process* $v = v_\xi \in L^2(\Omega \times [0,T]; \mathcal{H})$, *adapted to* $(\mathcal{F}_t)_{t\in[0,T]}$, *such that*

$$J_{0,T}\xi = \int_0^T J_{s,T} G v(s)\ ds,$$

and there exists a constant M, *independent of* ξ *and the initial value* x *of* $X(t;x)$, *such that*

$$\mathbb{E} \int_0^T |v(s)|^2\ ds \leq M|\xi|^2.$$

Proof: By [24], (B.26), there exists $u_1 \in L^2([0,T]; \mathcal{H})$ such that

$$S(T)(-\xi) + \int_0^T S(T-s) G u_1(s)\ ds = 0,$$

and

$$\int_0^T |u_1(s)|^2\ ds = |Q^{-1/2} S(T)\xi|^2 \leq ||Q^{-1/2} S(T)||^2 |\xi|^2. \tag{5.9}$$

Let $(\zeta(t))_{t\in[0,T]}$ be the solution of the pathwise inhomogeneous Cauchy problem

$$\begin{cases} \dot{\zeta}(t) = A\zeta(t) + dF(X(t;x)) J_{0,t}\xi + G u_1(t), \qquad t \geq 0, \\ \zeta(0) = 0. \end{cases} \tag{5.10}$$

Then

$$\begin{aligned} \zeta(T) &= \int_0^T S(T-s) dF(X(s,x)) J_{0,s}\xi\ ds + \int_0^T S(T-s) G u_1(s)\ ds \\ &= \int_0^T S(T-s) dF(X(s,x)) J_{0,s}\xi\ ds + S(T)\xi = J_{0,T}\xi. \end{aligned}$$

Define

$$u_2(t) := G^{-1} dF(X(t;x))[J_{0,t}\xi - \zeta(t)], \quad t \in [0,T], \text{ a.s.} \tag{5.11}$$

We see that $\zeta(t)$ also satisfies almost surely the inhomogeneous Cauchy problem

$$\begin{cases} \dot{\zeta}(t) = A\zeta(t) + dF(X(t;x))\zeta(t) + Gu_1(t) + Gu_2(t), & t \geq 0, \\ \zeta(0) = 0, \end{cases}$$

or, using variation of constants,

$$\zeta(t) = \int_0^t J_{s,t} Gv(s)\, ds,$$

where $v : \Omega \times [0,T] \to \mathcal{H}$ is defined by $v(t) := u_1(t) + u_2(t)$, $t \in [0,T]$. From (5.4), (5.9), (5.10) and (5.11) we see, using the Gronwall inequality, that $\mathbb{E} \int_0^t |v(s)|^2\, ds \leq M|\xi|^2$ for some $M > 0$ independent of ξ and x. □

The following corollary is a direct consequence of Lemma 5.7 and Lemma 5.8.

Corollary 5.9. *Under Hypothesis 5.5, and if $F : H \to H$ is Fréchet differentiable with uniformly bounded Fréchet derivative, then, for $\xi \in H$ and $v = v_\xi$ associated to ξ by Lemma 5.8, we have*

$$d_x X(T;x)\xi = J_{0,T}\xi = \int_0^T J_{s,T} Gv(s)\, ds = D^v X(T;x). \tag{5.12}$$

In (other) words: we have expressed the dependence of $X(T;x)$ on its initial condition x in terms of the dependence of $X(T;x)$ on the noise process W.

We can now give a short proof, as in [37], of the following theorem.

Theorem 5.10. *Under the conditions of Hypothesis 5.5 the transition semigroup associated to (5.1) is strong Feller at time T.*

PROOF: Suppose for the moment $\varphi \in C_b^1(H)$ and $F \in C_b^1(H;H)$. Let $(P(t))_{t\geq 0}$ denote the transition semigroup associated to (5.1). We have, using (5.12), the chain rule and integration by parts for the Malliavin derivative, that

$$\begin{aligned} |dP(T)\varphi(x)\xi| &= |\mathbb{E}\left[d\varphi(X(T;x))J_{0,T}\xi\right]| = |\mathbb{E}\left[d\varphi(X(T;x)D^v X(T;x)\right]| \\ &= |\mathbb{E}\left[D^v \varphi(X(T;x))\right]| = \left|\mathbb{E}\left[\varphi(X(T;x)) \int_0^T \langle v(s),\ dW(s)\rangle\right]\right| \\ &\leq ||\varphi||_\infty \left(\mathbb{E} \int_0^T |v(s)|^2\, ds\right)^{1/2}, \end{aligned}$$

where v is as described in Lemma 5.8, so that $\mathbb{E} \int_0^T |v(s)|^2\, ds \leq M^2|\xi|^2$ for some $M > 0$, independent of x and ξ. Hence

$$||dP(T)\varphi(x)||_{H^*} \leq M||\varphi||_\infty, \quad \text{for all } \varphi \in C_b^1(H), x \in H.$$

It follows that

$$|P(T)\varphi(x) - P(T)\varphi(y)| \leq M||\varphi||_\infty |x-y|_H, \quad \varphi \in C_b^1(H), x, y \in H.$$

We can extend this estimate to $\varphi \in B_b(H)$ and Lipschitz F by approximating φ by a sequence $(\varphi_n) \subset C_b^1(H)$, and F by a sequence $(F_n) \subset C_b^1(H;H)$ with $||dF_n||_\infty < [F]_{\mathrm{Lip}}$ (see the proof of [25], Theorem 7.1.1). □

5.3 Approximate controllability and irreducibility

Again we introduce a set of assumptions, similar to Hypothesis 5.3. Now we do not require null controllability of (A, G). These conditions which will be seen to be sufficient to prove the irreducibility of (5.1).

Hypothesis 5.11. *(i) H and $\mathcal{H}$ are Hilbert spaces;*

(ii) $(S(t))_{t\geq 0}$ is a strongly continuous semigroup acting on H with generator A;

(iii) W is a cylindrical Wiener process with RKHS $\mathcal{H}$;

(iv) $F : H \to V \subset \mathrm{im}\ (G)$, with V a closed subspace of H, is globally Lipschitz;

(v) $G \in L(\mathcal{H}; H)$ and a mapping $G^{-1} \in L(V, \mathcal{H})$ exists such that $GG^{-1} = I$ on V;

(vi) The pair (A, G) is approximately controllable at $T > 0$;

Proposition 5.12. *Suppose the assumptions of Hypothesis 5.11 hold. Then the system* (5.2)*, with $f = F$, is approximately controllable in time $T > 0$.*

Proof: Let $x, z \in H$ and $\varepsilon > 0$. Since (A, G) is approximately controllable, there exists a control $u_1 \in L^2([0,T];\mathcal{H})$ such that $\eta \in L^2([0,T];H)$ defined by

$$\eta(t) := S(t)x + \int_0^t S(t-s)Gu_1(s)\, ds \tag{5.13}$$

satisfies $|\eta(T) - z| < \varepsilon$.

For $0 \leq t \leq T$, choose

$$u_2(t) := -G^{-1}F(\eta(t)).$$

Then, for

$$u(t) := u_1(t) + u_2(t),$$

the solution of (5.2) is given by

$$\begin{aligned} y(t) &= S(t)x + \int_0^t S(t-s)F(y(s))\, ds + \int_0^t S(t-s)G(u_1(s) + u_2(s))\, ds \\ &= S(t)x + \int_0^t S(t-s)[F(y(s)) - F(\eta(s))]\, ds + \int_0^t S(t-s)Gu_1(s)\, ds. \end{aligned}$$

Suppose $(S(t))$ satisfies $||S(t)|| \leq M\mathrm{e}^{\omega t}$ for all $t \geq 0$ and some $M, \omega \geq 0$. Let

$$\zeta(t) := y(t) - \eta(t).$$

Then

$$\begin{aligned} |\mathrm{e}^{-\omega t}\zeta(t)| &= \left|\mathrm{e}^{-\omega t}\int_0^t S(t-s)[F(y(s)) - F(\eta(s))]\ ds\right| \\ &\leq \int_0^t M[F]_{\mathrm{Lip}}|\mathrm{e}^{-\omega s}\zeta(s)|\ ds, \quad t \geq 0, \end{aligned}$$

so that by Gronwall $\zeta \equiv 0$.

Hence $|y(T) - z| < \varepsilon$. □

Corollary 5.13. *Suppose the assumptions of Hypothesis 5.11 hold. Then the transition semigroup corresponding to the stochastic system (5.1) is irreducible in time T.*

Proof: This follows immediately from the approximate controllability proven in Proposition 5.12 and Theorem 7.4.1 in [25], which states that approximate controllability implies irreducibility. □

5.4 Examples

5.4.1 Stochastic delay differential equations

Consider, similar to [25], Section 10.2, a stochastic delay equation in $\mathbb{R}^d$ of the form

$$\begin{cases} dY(t) = \left(BY(t) + \sum_{i=1}^N B_i Y(t+\theta_i) + \varphi(Y(t), Y_t)\right)\ dt + \psi\ dW(t), \qquad t \geq 0 \\ Y(0) = x, \\ Y(\theta) = f(\theta), \quad \theta \in [-r, 0], \end{cases} \tag{5.14}$$

where $N \in \mathbb{N}$, $B, B_1, \ldots, B_N \in L(\mathbb{R}^d)$, $-r = \theta_1 < \theta_2 < \ldots < \theta_N < 0$, $\psi \in L(\mathbb{R}^m; \mathbb{R}^d)$, $(W(t))_{t\geq 0}$ an m-dimensional standard Brownian motion and the initial condition $x \in \mathbb{R}^d$. The segment process $(Y_t)_{t\geq 0}$ is defined by $Y_t(\theta) := Y(t+\theta)$ for $t \geq 0$, $-r \leq \theta \leq 0$, and $f \in L^2([-r, 0]; \mathbb{R}^d)$ is the initial segment. The nonlinear perturbation $\varphi : \mathbb{R}^d \times L^2([-r, 0]; \mathbb{R}^d) \to \mathbb{R}^d$ is assumed to be Lipschitz.

As explained in Section 3.2 we can cast this into the infinite dimensional framework (5.1) by choosing as Hilbert space $H := \mathbb{R}^d \times L^2([-r, 0]; \mathbb{R}^d)$, and letting the closed, densely defined operator A, described by

$$\begin{aligned} \mathfrak{D}(A) &= \left\{\begin{pmatrix} c \\ y \end{pmatrix} \in \mathbb{R}^d \times W^{1,2}([-r, 0]; \mathbb{R}^d) : y(0) = c\right\}, \\ A\begin{pmatrix} c \\ y \end{pmatrix} &:= \begin{pmatrix} Bc + \sum_{i=1}^N B_i y(\theta_i) \\ \dot{y} \end{pmatrix} \end{aligned}$$

denote the generator of a strongly continuous semigroup $(S(t))_{t\geq 0}$. (See e.g. [20], Section 2.4.)

As nonlinear perturbation $F : H \to H$ and noise factor $G \in L(\mathbb{R}^m; H)$ we take, respectively,

$$F\begin{pmatrix} c \\ y \end{pmatrix} := \begin{pmatrix} \varphi(c, y) \\ 0 \end{pmatrix}, \quad \text{and } G := \begin{pmatrix} \psi \\ 0 \end{pmatrix}.$$

For convenience, we recall the following result ([25], Theorem 10.2.3). See also [68] for the null controllability and [20], Theorem 4.2.10 for the approximate controllability.

Theorem 5.14. *The pair (A, G) is null controllable for all $t > r$ if and only if*

$$\operatorname{rank}\left[\lambda I - \sum_{i=1}^{N} \mathrm{e}^{\lambda\theta_i} B_i, \psi\right] = d \tag{5.15}$$

for all $\lambda \in \mathbb{C}$.

The pair (A, G) is approximately controllable for all $t > r$ if and only if

$$\operatorname{rank}\left[\lambda I - B - \sum_{i=1}^{N} \mathrm{e}^{\lambda\theta_i} B_i, \psi\right] = d, \quad \operatorname{rank}[B_1, \psi] = d \tag{5.16}$$

for all $\lambda \in \mathbb{C}$.

Remark 5.15. The above theorem is partly based on [68]. In this paper null controllability after some time $t > 0$ is established; however from the proof in this paper it is not clear whether null controllability holds for all $t > r$. This has no significant consequence since, without loss of generality, we may take $r > 0$ large enough so that we indeed have null controllability for all $t > r$.

We can now state the main result of this section.

Theorem 5.16. *Suppose conditions (5.15) and (5.16) are satisfied. Let $\widetilde{V}$ be a linear subspace of $\mathbb{R}^d$ such that $\overline{\varphi(H)} \subset \widetilde{V}$. Suppose that a mapping $\psi^{-1} \in L(\widetilde{V}; \mathbb{R}^m)$ of ψ exists, i.e. $\psi\psi^{-1}v = v$ for $v \in \widetilde{V}$.*

Then there exists at most one invariant probability measure for (5.14) on the state space H, and if an invariant probability measure exists, it is strongly mixing.

PROOF: Define $G^{-1} \in L(\widetilde{V} \times \{0\}; \mathbb{R}^m)$ by $G^{-1}\begin{pmatrix} v \\ 0 \end{pmatrix} := \psi^{-1}v$. All the conditions of Hypothesis 5.3 are satisfied (with $T > r$ and $V = \widetilde{V} \times \{0\} \subset \mathbb{R}^d \times L^2([-r, 0]; \mathbb{R}^d)$), and by Theorem 5.4 we may deduce the uniqueness of an invariant probability measure and the strong mixing property of such a measure. □

Note that the conditions of Theorem 5.16 are not necessarily very restrictive:

Corollary 5.17. *Suppose that $m \geq d$ and $\psi \in L(\mathbb{R}^m; \mathbb{R}^d)$ is surjective.*

Then there exists at most one invariant probability measure for (5.14) on the state space H, and if it exists, it is strongly mixing.

PROOF: Define the pseudoinverse ψ^{-1} by letting $\psi^{-1}v$ denote the element w of minimal norm in $\mathbb{R}^m$ such that $\psi w = v$. Then $\psi^{-1} \in L(\mathbb{R}^d; \mathbb{R}^m)$ is a linear operator. Since $m \geq d$ and ψ is surjective, we find that rank $\psi = d$ and hence (5.15) and (5.16) hold. The result follows now from Theorem 5.16, of which all conditions are satisfied (with $\widetilde{V} = \mathbb{R}^d$). □

For convenience we combine our result with a result of Chapter 4 on the existence of invariant probability measures.

Corollary 5.18. *Suppose the solutions of (5.14) are bounded in probability on the state space H, and the conditions of Theorem 5.16 hold. Then there exists a unique, strongly mixing invariant probability measure for (5.14) on H.*

PROOF: The existence of an invariant measure under these conditions is proven in Theorem 4.3. The uniqueness follows from Theorem 5.16. □

5.4.2 Stochastic reaction-diffusion recurrent neural networks

In [54] the following stochastic partial differential equation in m dimensions with delay and noise is considered as an example of so-called *recurrent neural networks*.

$$
\begin{aligned}
dy_i(t) = \sum_{k=1}^{m} \frac{\partial}{\partial \xi_k}\left(D_i \frac{\partial y_i}{\partial \xi_k}\right)\, dt + \Bigg[& -c_i h_i(y_i(t,\xi)) + \sum_{j=1}^{n} a_{ij} f_j(y_j(t,\xi)) \\
& + \sum_{j=1}^{n} b_{ij} \int_{-\infty}^{t} \kappa_{ij}(t-s) g_j(y_j(s,\xi))\, ds + J_i \Bigg]\, dt \\
& + \sum_{l=1}^{\infty} \sigma_{il}(y_i(t,\xi))\, dw_{il}(t).
\end{aligned}
$$

We consider the following variant for n neurons in one dimension:

$$
\begin{aligned}
dy_i(t,\xi) = \Delta y_i(t,\xi)\, dt + \Bigg[& -c_i h_i(y_i(t,\xi)) + \sum_{j=1}^{n} a_{ij} f_j(y_j(t,\xi)) \\
& + \sum_{j=1}^{n} b_{ij} \int_{t-1}^{t} \kappa_{ij}(t-s) g_j(y_j(s,\xi))\, ds + J_i \Bigg]\, dt + (\Psi_i\, dW_i(t))(\xi),
\end{aligned}
\tag{5.17}
$$

where

(i) Δ denotes the one-dimensional Laplacian $\frac{d^2}{d\xi^2}$ on $[0,\pi]$;

(ii) c_i, a_{ij}, b_{ij}, J_i are constants for $i = 1,\dots,n$, $j = 1,\dots,n$;

(iii) h_i, f_j and g_j are Lipschitz functions $\mathbb{R} \to \mathbb{R}$ for $i,j = 1,\dots,n$;

(iv) $\kappa_{ij} \in L^2([0,1])$ for $i,j = 1,\dots,n$;

(v) $\Psi_i \in L(\mathcal{H}_i, L^2([0,\pi]))$, where $\mathcal{H}_i$ is the RKHS of the cylindrical Wiener process W_i, $i = 1,\dots,n$.

Let $X = L^2([0,\pi]) \times \dots \times L^2([0,\pi])$ and let the delay semigroup $(S(t))$ on $\mathcal{E}^2 = X \times L^2([-1,0];X)$ be generated by

$$A = \begin{bmatrix} B & 0 \\ 0 & \frac{d}{d\sigma} \end{bmatrix} \tag{5.18}$$

with

$$B = \begin{bmatrix} \Delta & 0 & \dots & 0 \\ 0 & \Delta & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \Delta \end{bmatrix}.$$

Note that there is no dependence on the past in the generator of the delay semigroup. This will be in our advantage later on.

Denote typical elements of $\mathcal{E}^2$ by $z = \begin{pmatrix} u_1 & \dots & u_n \\ v_1 & \dots & v_n \end{pmatrix}$ with $u_i \in L^2([0,\pi])$ and $v_i \in L^2([-1,0];L^2([0,\pi]))$, $i = 1,\dots,n$. Let $\varphi_i : X \times L^2([-1,0];X) \to L^2([0,\pi])$ be given by

$$\begin{aligned} &\varphi_i(u_1,\dots,u_n,v_1,\dots,v_n)[\xi] \\ &:= -c_i h_i(u_i(\xi)) + \sum_{j=1}^n a_{ij} f_j(u_j(\xi)) + \sum_{j=1}^n b_{ij} \int_{-1}^0 \kappa_{ij}(-s) g_j(v_j(s,\xi))\, ds + J_i, \end{aligned}$$

and define $F : X \times L^2([-1,0];X) \to X \times L^2([-1,0];X)$ and $G \in L(\mathcal{H}_1 \times \dots \times \mathcal{H}_n; X \times L^2([-1,0];X))$ by

$$F(z) := \begin{pmatrix} \varphi_1(z) & \dots & \varphi_n(z) \\ 0 & \dots & 0 \end{pmatrix}, \quad z \in \mathcal{E}^2, \quad \text{and} \tag{5.19}$$

$$G\begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} := \begin{pmatrix} \Psi_1 w_1 & \dots & \Psi_n w_n \\ 0 & \dots & 0 \end{pmatrix}, \quad (w_1,\dots,w_n) \in \mathcal{H}_1 \times \dots \times \mathcal{H}_n. \tag{5.20}$$

Now (5.17) can be written in the form (5.1).

Proposition 5.19. *There exists a unique solution to* (5.1)*, with A, F and G as given in* (5.18)*,* (5.19) *and* (5.20)*.*

PROOF: This can be shown in a similar way as in the first example of Section 3.4.1, only here we consider a cartesian product of function spaces $L^2([0,\pi])$. $\square$

We will now establish a sufficient condition for (A, G) to be eventually null controllable.

Theorem 5.20. *Suppose $\Psi_i \in L(\mathcal{H}_i; L^2([0,\pi]))$ has a bounded inverse for all $i = 1, \ldots, n$. Then (A, G) is null controllable for $t > 1$.*

PROOF: Since A and G are 'in diagonal form' it suffices to consider the case where $n = 1$, so let $\Psi \in L(\mathcal{H}; X)$ be invertible with $\Psi^{-1} \in L(X; \mathcal{H})$. It is sufficient to establish the null controllability of (Δ, Ψ) on $X = L^2([0,\pi])$. Indeed, if (Δ, Ψ) is null controllable, then for any $t > 1$ and $x \in X \times L^2([-1,0]; X)$ we may find a control which steers the first component of x to 0 in time $t-1$. By setting the control equal to zero after time $t-1$ the translation effect of the delay semigroup then ensures that x is steered to 0 in $X \times L^2([-1,0]; X)$ in time t.

It remains to establish the null controllability of (Δ, Ψ) on $L^2([0,\pi])$. By [20], condition (4.12), this is equivalent to the the existence of a $\gamma(t) > 0$ for all $t > 0$ such that

$$\int_0^t ||\Psi^* T^*(t-s)z||^2_{\mathcal{H}}\ ds \geq \gamma(t)||T^*(t)z||^2_X \tag{5.21}$$

for all $z \in X$, where $(T(t))_{t\geq 0}$ is the semigroup generated by the Laplacian. Let $(e_n)_{n\in\mathbb{N}\cup\{0\}}$ be the orthonormal base of eigenvectors of the Laplacian with Neumann boundary conditions on $X = L^2([0,\pi])$, so

$$\begin{cases} e_0(\xi) = \frac{1}{\sqrt{\pi}}, \\ e_n(\xi) = \sqrt{\frac{2}{\pi}}\cos(n\xi), \qquad n \in \mathbb{N}, \end{cases}$$

and for $z \in X$ write $z_n := \langle z, e_n\rangle_X$ for $n \in \mathbb{N}\cup\{0\}$. We have, using selfadjointness of $(T(t))_{t\geq 0}$,

$$\begin{aligned}
\int_0^t ||\Psi^* T^*(t-s)z||^2_{\mathcal{H}}\ ds &\geq \frac{1}{||\Psi^{-1}||^2}\int_0^t ||T(t-s)z||^2_X\ ds \\
&= \frac{1}{||\Psi^{-1}||^2}\int_0^t \left|\left|\sum_{n=0}^\infty T(t-s)z_n e_n\right|\right|^2_X\ ds \\
&= \frac{1}{||\Psi^{-1}||^2}\int_0^t \left[z_0^2 + \sum_{n=1}^\infty \mathrm{e}^{-2n^2 s} z_n^2\right] ds \\
&= \frac{1}{||\Psi^{-1}||^2}\left[t z_0^2 + \sum_{n=1}^\infty \frac{1}{2n^2}(1-\mathrm{e}^{-2n^2 t})z_n^2\right],
\end{aligned}$$

which should be compared to

$$||T^*(t)z||_X^2 = z_0^2 + \sum_{n=1}^{\infty} \mathrm{e}^{-2n^2 t} z_n^2.$$

Using the basic inequality

$$\frac{1}{a}\left(1 - \mathrm{e}^{-at}\right) \geq t\mathrm{e}^{-at}, \quad a, t > 0,$$

we find that (5.21) holds for

$$\gamma(t) = \frac{t}{||\Psi^{-1}||^2},$$

which establishes the null controllability of (Δ, Ψ). □

Corollary 5.21. *The transition semigroup corresponding to* (5.1) *with A, F and G as given in* (5.18), (5.19) *and* (5.20), *is eventually strong Feller.*

PROOF: This is an immediate corollary of the previous theorem and Theorem 5.10, using the invertibility of Ψ_i, $i = 1, \ldots, n$. □

5.5 Asymptotic strong Feller property

The strong Feller which we discussed earlier (Section 5.1 and Section 5.2) tells something about the smoothing properties of the Markov transition semigroup. However, as Hairer and Mattingly described in [37], also contractive properties can be used to establish uniqueness of solutions. Therefore they introduced the *asymptotic strong Feller property* which is a generalization of the strong Feller property.

We will refrain from giving a precise definition (which is rather intricate and would need a long introduction). Instead, we quote a sufficient condition ([37], Proposition 3.12) for a Markov semigroup to possess the asymptotic strong Feller property.

Proposition 5.22. *Let $(t_n)_{n\in\mathbb{N}}$ and $(\delta_n)_{n\in\mathbb{N}}$ be two sequences of positive real numbers with $(t_n)_{n\in\mathbb{N}}$ nondecreasing and $(\delta_n)_{n\in\mathbb{N}}$ converging to zero. A transition semigroup $(P(t))_{t\geq 0}$ of a Markov process with values in a Hilbert space H is asymptotically strong Feller if, for all $\varphi : H \to \mathbb{R}$ with $||\varphi||_\infty$ and $||d\varphi||_\infty$ finite,*

$$||d(P(t_n)\varphi)(x)|| \leq C(|x|)\left(||\varphi||_\infty + \delta_n ||d\varphi||_\infty\right)$$

for all $n \in \mathbb{N}$ and $x \in H$, and where $d\varphi$ denotes the Fréchet derivative of φ and $C : [0, \infty) \to \mathbb{R}$ is a fixed nondecreasing function.

In order to state the promised ergodic property following from the asymptotic strong Feller property, we need the notion of support of a measure.

Definition 5.23. *For any measure μ on a topological space X the* support $\operatorname{supp}\mu$ *is defined as the smallest closed set $F \subset X$ with $\mu(X \backslash F) = 0$.*

Here we quote the result ([37], Theorem 3.16) which explains the importance of the asymptotic strong Feller property.

Theorem 5.24. *Let $(P(t))_{t\geq 0}$ be a Markov semigroup and let μ and ν be two distinct ergodic invariant probability measures for $(P(t))_{t\geq 0}$. If $(P(t))_{t\geq 0}$ is asymptotically strong Feller at x, then $x \notin \operatorname{supp}\mu \cap \operatorname{supp}\nu$.*

As noted in [37], this theorem has the following corollary.

Corollary 5.25. *If $(P(t))_{t\geq 0}$ is an asymptotically strong Feller semigroup and there exists a point x such that $x \notin \operatorname{supp}\mu$ for every invariant probability measure μ of $(P(t))_{t\geq 0}$ then there exists at most one invariant probability measure for $(P(t))_{t\geq 0}$.*

The following proposition gives a sufficient condition for certain linear stochastic evolutions to possess the asymptotic strong Feller property.

Proposition 5.26. *Let H be a Hilbert space with closed linear subspaces V and W (not necessarily orthogonal) such that $H = V \oplus W$. Let $\pi : H \to V$ denote the projection on V along W. Let $G \in L_{\mathrm{HS}}(\mathcal{H}; H)$. Let $(S(t))_{t\geq 0}$ be a strongly continuous semigroup on H such that the restriction of S to W is asymptotically stable, i.e. $|| \, S(t)|_W \, || \to 0$, $t \to \infty$. Furthermore suppose that $(\, S(t)|_V \,, \pi G)$ is null controllable at time $T > 0$, i.e. for any $v \in V$ there exists a function $u \in L^2([0,T]; \mathcal{H})$ such that*

$$S(T)v + \int_0^T S(T-s)\pi G u(s) \; ds = 0.$$

Then the Markov transition semigroup $(P(t))_{t\geq 0}$ for

$$dX = AX \; dt + G \; dW(t) \tag{5.22}$$

posesses the asymptotic strong Feller property.

Proof: Let $z \in H$ and $t \geq T$. Let $u \in L^2([0,t]; \mathcal{H})$ such that

$$S(t)\pi z + \int_0^t S(t-s)\pi G u(s) \; ds = 0$$

and let $\varphi : H \to \mathbb{R}$ with $||\varphi||_\infty \vee ||D\varphi||_\infty < \infty$ and $v \in V$.

For $m \in H$ and $Q \in L_1(H)$ let $N_{m,Q}$ denote the normal distribution on H with mean m and covariance operator Q. Furthermore let $N_Q := N_{0,Q}$ and let $\rho_t(m, \cdot)$ denote the Radon-Nikodym derivative

$$\rho_t(m,y) := \frac{dN_{m,Q_t}}{dN_{Q_t}}(y), \quad y \in H,$$

for $m \in Q_t^{1/2}(H)$ (see [26], Theorem 1.3.6 or [21], Theorem 2.8).

Then,

$$\begin{aligned}
D_x P(t)\varphi(x)v &= \lim_{h\to 0}\frac{1}{h}\int_H \{\varphi(y+S(t)(x+hv)) - \varphi(y+S(t)x)\}\ N_{Q_t}(dy) \\
&= \lim_{h\to 0}\frac{1}{h}\left\{\int_H \varphi(y+S(t)x)\ N_{hS(t)v,Q_t}(dy)\right. \\
&\qquad\qquad \left. - \int_H \varphi(y+S(t)x)\ N_{Q_t}(dy)\right\} \\
&= \lim_{h\to 0}\frac{1}{h}\int_H \varphi(y+S(t)x)\,(\rho_t(hS(t)v,y)-1)\ N_{Q_t}(dy) \\
&= \int_H \varphi(y+S(t)x)\ \frac{d}{dh}\left(\rho_t(hS(t)v,y)-1\right)\Big|_{h=0}\ N_{Q_t}(dy) \\
&= \int_H \varphi(y+S(t)x)\left\langle Q_t^{-1/2}y, \Gamma(t)v\right\rangle\ N_{Q_t}(dy),
\end{aligned}$$

and

$$\left|\int_H \varphi(y+S(t)x)\left\langle Q_t^{-1/2}y,\Gamma(t)v\right\rangle\ N_{Q_t}(dy)\right| \le ||\varphi||_\infty |\Gamma(t)v|_H,$$

with, by the Cameron-Martin formula (see [21], Theorem 1.3.6),

$$\Gamma(t)v := Q_t^{-1/2}S(t)v, \quad \text{for} \quad t \ge 0, v \in H.$$

Note that $\Gamma(t)$ is defined by null controllability of $(\left.S(t)\right|_V, \pi G)$, and that

$$||\Gamma(t)||_{L(V;H)} \le ||\Gamma(T)||_{L(V;H)}, \quad \text{for } t \ge T,$$

since $||\Gamma(t)v||^2$ denotes the 'minimal energy' required to steer v to 0 in time t and is hence decreasing in t; see [21], Section 8.3.1.

Now for $w \in W$, we have

$$\begin{aligned}
|D_x P(t)\varphi(x)w| &= \left|\left(D_x\int_H \varphi(y+S(t)x)\ N_{Q_t}(dy)\right)w\right| \\
&= \left|\int_H D\varphi(y+S(t)x)\ N_{Q_t}(dy)S(t)w\right| \\
&\le ||D\varphi||_\infty |S(t)w|_H.
\end{aligned}$$

Now for $z = v + w \in H$ with $v \in V$ and $w \in W$, we find that

$$|D_x P(t)\varphi(x)z| \le ||\Gamma(T)||_{L(V;H)}||\varphi||_\infty |v| + ||D\varphi||_\infty |S(t)w|_H.$$

We may now apply the previously stated sufficient condition for the asymptotic strong Feller property (Proposition 5.26). □

Note that the condition of the above proposition is closely related to that of stabilizability (see also Section 6.3 of this thesis).

5.5.1 Example: eventually compact semigroups

Recall that a strongly continuous semigroup $(T(t))_{t\geq 0}$ with infinitesemal generator A is called eventually compact if there exists a $t_0 > 0$ such that $T(t_0)$ is a compact operator. Eventually compact semigroups allow a decomposition into two closed subspaces, of which one is finite dimensional, as we discuss below. Therefore, to establish the asymptotic strong Feller property for systems driven by an eventually compact semigroup, it remains only to check the finite dimensional controllability of the non-stable part. The delay semigroup is an important example of an eventually compact semigroup, see Section 3.5.

In Section IV.2 of [29] the spectral decomposition of eventually compact operators is discussed. We summarize the relevant definitions and statements below.

Let $L : \mathfrak{D}(L) \to X$ be a linear operator. The *point spectrum* $\sigma_p(L)$ is the set of those $\lambda \in \mathbb{C}$ for which $\ker(\lambda I - L) \neq \{0\}$. We call $\lambda \in \sigma_p(L)$ an *eigenvalue* of L. The nullspace $\ker(\lambda I - L)$ is called the *eigenspace*. The *generalized eigenspace* $\mathcal{M}_\lambda = \mathcal{M}_\lambda(L)$ is the smallest closed linear subspace that contains all $\ker\big((\lambda I - L)^j\big)$ for $j = 1, 2, \ldots$.

Theorem 5.27 (Spectral resolution of eventually compact semigroups)**.** *Let $(T(t))_{t\geq 0}$ be an eventually compact semigroup. Let $\beta \in \mathbb{R}$. The set*

$$\Lambda = \Lambda(\beta) := \{\lambda \in \sigma(A) : \mathrm{Re} \quad \lambda > \beta\}$$

is a finite set of isolated eigenvalues of A. There exists a closed subspace $\mathcal{R}_\Lambda$ such that

$$X = \mathcal{M}_\Lambda \oplus \mathcal{R}_\Lambda,$$

where

$$\mathcal{M}_\Lambda = \oplus_{\lambda\in\Lambda}\mathcal{M}_\lambda(A)$$

has finite dimension.

Let

$$P_\Lambda := \sum_{\lambda\in\Lambda} \mathrm{Res}_{\ z=\lambda} R(z, A) = \sum_{\lambda\in\Lambda} \frac{1}{2\pi\mathrm{i}} \int_{\Gamma_\lambda} R(z, A)\ dz,$$

with Γ_λ a small circle in C such that λ is the only eigenvalue of A inside Γ_λ. Then P_Λ is called the spectral projection *onto $\mathcal{M}_\Lambda$ along $\mathcal{R}_\Lambda$, and there exist positive constants K and δ such that*

$$||T(t)(I - P_\Lambda)|| \leq K\mathrm{e}^{(\beta+\delta)t}||I - P_\Lambda||, \quad t \geq 0,$$

Corollary 5.28. *Suppose A generates a strongly continuous eventually compact semigroup in H, $G \in L(\mathcal{H}; H)$, $\beta < 0$ and $(A|_{M_{\Lambda(\beta)}}, P_{\Lambda(\beta)}G)$ is controllable. Then the evolution described by (5.22) is asymptotically strong Feller.*

5.6 Notes and remarks

In [58] and [60] an overview is given of results on uniqueness of invariant probability measures and on strong Feller diffusions, respectively. In [59] the same authors present similar results for stochastic delay equations as those found in this chapter. Their conditions on the evolution process are probabilistic in nature, whereas our conditions are of an operator theoretic kind.

Very recently the uniqueness of invariant probability measure for general stochastic delay equations with multiplicative noise was established in [36] .

In [70] uniqueness of an invariant probability measure was established for nondegenerate diffusions in Hilbert spaces, and in [19] for degenerate diffusions. However, in the latter, only the immediate strong Feller property was established which is too strong for our purposes: the delay semigroup can never be immediately strong Feller.

In [33] and [34] the immediate strong Feller property and irreducibility are proven for (possibly degenerate) diffusions, by applying Malliavin calculus. Their result does not apply to stochastic delay equations since these can only be eventually strong Feller.

Uniqueness of invariant probability measure in Banach spaces is discussed in [87]. See also references mentioned in Section 4.4

Based on the contents of this chapter, a paper is being prepared for publication ([10]).

CHAPTER

SIX

Stability

In this chapter we consider the linear SDE

$$dX(t) = AX(t)\ dt + \sum_{i=1}^{n} B_i X(t)\ dW(t)$$

or in mild form

$$X(t) = S(t)x + \sum_{i=1}^{n} \int_0^t S(t-s) B_i X(s)\ dW(s). \tag{6.1}$$

with $(S(t))_{t\geq 0}$ a strongly continuous semigroup in a Hilbert space H with infinitesemal generator A, $B_i \in L(H)$, $i = 1, \ldots, n$, $(W_i)_{i=1}^n$ independent standard Brownian motions in $\mathbb{R}$ and $x \in H$. By Theorem 2.27 there exists a unique solution $(X(t;x))_{t\geq 0}$ to (6.1) for any choice of A, $(B_i)_{i=1}^k$ and x.

In this paper we are interested in the stability properties of the solution of (6.1). More specifically, we want to estimate the *pathwise Lyapunov exponent*, defined (see [5]) as

$$\lambda := \lim_{t\to\infty} \frac{1}{t} \log |X(t;x)|, \quad \text{a.s.} \tag{6.2}$$

An inspiration was the paper [2] by John Appleby and Xuerong Mao concerning stochastic stability of functional differential equations. Treating functional differential equations from a finite dimensional perspective they managed to find conditions for pathwise stability. Our goal was to obtain similar stability results by an infinite dimensional approach (i.e. by using the infinite dimensional state space for the delay equation).

In Section 6.1 we will discuss some preliminary results. In Section 6.2 we will briefly discuss the history of the pathwise stochastic stability problem in finite dimensions. In Section 6.3 we will establish a necessary condition on the drift and noise parts of the SDE in order for the solutions to be stochastically stable. In Section 6.4 we will discuss the strong law of large numbers and the approximation of solutions of SDEs in Hilbert space using Yosida approximations, both of which are important tools for the remainder of this chapter. In Section 6.5 an estimate on the pathwise Lyapunov exponent for the case of nondegenerate noise is discussed. Finally in Section 6.6 we discuss a technique to find an estimate on the pathwise Lyapunov exponent for degenerate noise by means of an operator inequality.

6.1 Preliminary results

6.1.1 Deterministic systems

Let A be a closed linear operator and let $\rho(A)$ denote the *resolvent set* of A:

$$\rho(A) = \left\{\lambda \in \mathbb{C} : (\lambda - A)^{-1} \text{ exists and is bounded}\right\}.$$

Define the *spectrum* $\sigma(A)$ of A by $\sigma(A) = \mathbb{C}\backslash\rho(A)$. Let $\mathfrak{s}(A)$ denote the *spectral bound* of A, and $\mathfrak{r}(B)$ the *spectral radius* of B, i.e.

$$\mathfrak{s}(A) := \sup\{\text{Re } \lambda : \lambda \in \sigma(A)\} \quad \text{and} \quad \mathfrak{r}(B) := \sup\{|\lambda| : \lambda \in \sigma(B)\}.$$

Furthermore let $\omega_0(A)$ denote the *growth bound* of A, i.e.

$$\omega_0(A) := \inf\{\omega \in \mathbb{R} : \exists_{M\geq 1} \text{ s.t. } ||\exp(At)|| \leq M\mathrm{e}^{\omega t} \text{ for all } t \geq 0\}.$$

We have the following relation between $s(A)$, $\omega_0(A)$ and $r(S(t))$:

$$\mathfrak{s}(A) \leq \omega_0(A) = \frac{1}{t}\log\ \mathfrak{r}\left(S(t)\right), \quad \text{for all } t \geq 0. \tag{6.3}$$

See [32], Proposition IV.2.2. If $(S(t))_{t\geq 0}$ is eventually norm continuos then by [32], Theorem IV.3.11, the inequality in (6.3) becomes an equality: $s(A) = \omega_0(A)$.

6.1.2 Commutative case

As an appetizer, consider the particular case of (6.1) where A and all B_i, $i = 1, \ldots, k$ commute. Let $(T_0)_{t\geq 0}$ be the strongly continuous semigroup generated by $A - \frac{1}{2}\sum_{i=1}^k B_i^2$, and let $(T_i)_{t\geq 0}$ be the uniformly continuous groups $(T_i(t))_{t\geq 0}$ generated by B_i, $i = 1, \ldots, k$. Then the solution is given by

$$X(t; x) = T_0(t)\prod_{i=1}^k T_i(W_i(t))x. \quad t \geq 0, x \in H.$$

Suppose $||T_i(t)|| \leq M_i \exp(\omega_0(B_i)t)$, $i = 1, \ldots, k$. Then

$$\frac{1}{t}\log|X(t;x)| \leq \frac{1}{t}\log|x| + \frac{1}{t}\log||T_0(t)|| + \frac{1}{t}\sum_{i=1}^{k}\log||T_i(W_i(t))||.$$

Now using the strong law of large numbers for martingales (see Theorem 6.11)

$$\limsup_{t\to\infty}\frac{1}{t}\log||T_i(W_i(t))|| \leq \limsup_{t\to\infty}\frac{\log M_i}{t} + \frac{\omega_0(B_i)|W_i(t)|}{t} = 0 \quad \text{a.s.},$$

and furthermore

$$\limsup_{t\to\infty}\frac{1}{t}\log||T_0(t)|| = \omega_0\left(A - \tfrac{1}{2}\sum_{i=1}^{k}B_i^2\right).$$

Hence we find

$$\limsup_{t\to\infty}\frac{1}{t}\log|x(t)| \leq \omega_0\left(A - \tfrac{1}{2}\sum_{i=1}^{k}B_i^2\right), \quad \text{a.s.} \tag{6.4}$$

This may be compared to a result by A. Kwiecińska, [49], which seems to be slightly less sharp and less general.

6.1.3 Interpretation of the pathwise Lyapunov exponent

The expression for the Lyapunov exponent given in (6.2) implies that there exists a random variable M such that

$$|X(t;x)| \leq M\mathrm{e}^{\lambda t} \quad \text{for all } t \geq 0, \text{ a.s.}$$

It should be noted that if we can estimate $\inf_{t\geq 0}\mathbb{E}[|X(t;x)|^p] \geq k^p$ for some $p > 0$, then

$$k^p \leq \mathbb{E}|X(t;x)|^p \leq \mathbb{E}[M^p\mathrm{e}^{\lambda pt}],$$

so that

$$\mathbb{E}M^p \geq k^p\mathrm{e}^{-\lambda pt} \quad \text{for all } t \geq 0,$$

and in particular $\mathbb{E}M^p = \infty$ if $\lambda < 0$.

So in this sense a negative Lyapunov exponent does not guarantee a fast convergence to zero. Examples of this behaviour are easy to construct: in one dimension take $A = 0$ and $B_1 = b \in \mathbb{R}$; then $\mathbb{E}|X(t;x)|^2 \geq \mathbb{E}|x|^2$, $t \geq 0$, but $\lambda = -\frac{1}{2}b^2$.

6.2 History of the problem in finite dimensions

In this section we present a short history of establishing the pathwise Lyapunov exponent for solutions of linear SDEs. This is for illustrative purposes only, and we do not claim any completeness. The methods described in this section will not be used later on.

An early reference to the problem is found in [50], where it is postulated that if the solution of

$$\dot{x}(t) = Ax(t), \quad t \geq 0$$

is unstable, then the solution of the Stratonovich SDE

$$dx(t) = Ax(t) + \sum_{i=1}^{k} B_i x(t) \circ dW_i(t),$$

is never (pathwise) stable.

Already in 1967, Khas'minskii [46] shows that this postulate does not hold by studying (6.1) in spherical coordinates. His approach is as follows.

First we write (6.1) in Stratonovich form

$$dx(t) = \widetilde{A}x(t)\ dt + \sum_{i=1}^{k} B_i x(t) \circ dW_i(t), \tag{6.5}$$

with $\widetilde{A} = A - \frac{1}{2}\sum_{i=1}^{k} B_i^2$.

Write $y(t) := x(t)/|x(t)|$ and $\lambda(t) := \log|x(t)|$, $t \geq 0$. Then y is a process on S^{n-1}, the unit sphere of dimension $n-1$. Using Itô's formula it can be calculated that y satisfies

$$\begin{cases} dy(t) = h(\widetilde{A}, y(t))\ dt + \sum_{i=1}^{k} h(B_i, y(t)) \circ dW_i(t), \quad t \geq 0, \\ y(0) = y_0 := \frac{x_0}{|x_0|}, \end{cases} \tag{6.6}$$

where

$$h(C, z) := (C - q(C, z)I)z, \quad q(C, z) := z^T C z, \quad z \in S^{n-1}, C \in \mathbb{R}^{n \times n}.$$

Note that $h(C, z)^T z = 0$ for $z \in S^{n-1}$ so that indeed $h(C, z)$ is a vector in the tangent space $T_z S^{n-1}$ and that the process $(y(t))_{t \geq 0}$ is autonomous, as can be seen from (6.6). By compactness of S^{n-1} at least one invariant measure μ exists for y.

Khas'minskii then assumes a strong non-singularity condition on the (B_i),

$$\sum_{i=1}^{k} (B_i x)(B_i x)^T \text{ is positive definite for all } x \in \mathbb{R}^n,\ x \neq 0. \tag{6.7}$$

Due to this condition, $\mathbb{P}_{y_0}(y(t) \in U) > 0$ for all open $U \subset S^{n-1}$ and $(y(t))_{t\geq 0}$ is strong Feller. By an earlier theorem of Khas'minskii ([45]), μ is therefore the unique invariant measure for $(y(t))_{t\geq 0}$ on S^{n-1}, and hence it is ergodic.

By Itô's formula, the process $(\lambda(t))_{t\geq 0}$ can be shown to satisfy

$$\lambda(t) = \lambda(0) + \int_0^t \Phi(y(s))\ ds + \sum_{i=1}^k \int_0^t \langle B_i y(s), y(s)\rangle\ dW_i(s) \quad \text{a.s.}, \tag{6.8}$$

with

$$\Phi(z) := \langle Az, z\rangle + \tfrac{1}{2}\sum_{i=1}^k ||B_i z||^2 - \sum_{i=1}^k \langle B_i z, z\rangle^2, \quad z \in S^{n-1}. \tag{6.9}$$

Now using the strong law of large numbers for martingales (Theorem 6.11),

$$\lim_{t\to\infty} \frac{1}{t}\sum_{i=1}^k \int_0^t \langle B_i y(s), y(s)\rangle\ dW_i(s) = 0 \quad \text{a.s.},$$

and by ergodicity of μ,

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t \Phi(y(s))\ ds = \int_{S^{n-1}} \Phi(z)\ d\mu(z) \quad \text{a.s.}$$

We may conclude that

$$\lim_{t\to\infty} \frac{1}{t}\log|x(t)| = \int_{S^{n-1}} \Phi(z)\ d\mu(z) \quad \text{a.s.} \tag{6.10}$$

As stated above, (6.7) is unnecessarily strong for establishing the uniqueness of the invariant measure μ on S^{n-1}. A better understanding of the structure of ergodic invariant measures on manifolds (e.g. S^{n-1}) is provided by [48].

In [55], Mao provides a new way of estimating the Lyapunov exponent. In the linear case this boils down to requiring that

$$\langle Az, z\rangle \leq \alpha \quad \text{and} \quad \tfrac{1}{2}\sum_{i=1}^k ||B_i z||^2 - \sum_{i=1}^k \langle B_i z, z\rangle^2 \leq \beta, \quad z \in S^{n-1} \tag{6.11}$$

so that $\Phi(z) \leq \alpha + \beta$, $z \in S^{n-1}$. Therefore (6.8) shows that

$$\limsup_{t\to\infty} \frac{1}{t}\log|x(t)| \leq \alpha + \beta \quad \text{a.s.},$$

without the need to establish an invariant measure on S^{n-1}.

This approach can be extended using a Lyapunov function, see [56], Theorem 4.3.3.

6.3 Necessary condition for pathwise stability

Throughout this section, let H and U be Hilbert spaces, $(T(t))_{t\geq 0}$ a strongly continuous semigroup on H with infinitesemal generator $A : \mathfrak{D}(A) \to H$, $B \in L(U;H)$, $G \in L(H)$ and $F \in L(H;U)$. Furthermore $(M(t))_{t\geq 0}$ is a one-dimensional square integrable martingale.

In this section we will compare stability properties of the stochastic differential equation

$$dX(t) = AX(t)\ dt + BFX(t)\ dM(t)$$

to stabilizability of the linear control system

$$x'(t) = Ax(t) + Bu(t).$$

First we introduce some necessary notions of stability for stochastic processes.

Definition 6.1. *Let $(X(t))_{t\geq 0}$ be a stochastic process in H. The process $(X(t))_{t\geq 0}$ is said to be* almost surely asymptotically stable *or* pathwise asymptotically stable *if*

$$\lim_{t\to\infty} X(t) = 0, \quad \textit{almost surely},$$

and $(X(t))_{t\geq 0}$ is said to be almost surely exponentially stable *or* pathwise exponentially stable *if*

$$\limsup_{t\to\infty} \frac{1}{t} \log |X(t)| < 0, \quad \textit{almost surely.}$$

Suppose $(X(t;x))_{t\geq 0}$ is the unique solution of

$$dX(t) = AX(t)\ dt + GX(t)\ dM(t), \quad X(0) = x \in H,$$

and that $(X(t;x))_{t\geq 0}$ is pathwise asymptotically stable for all $x \in H$. Then the pair (A, G) is called stochastically stable.

We will compare the stability properties above to the following deterministic stability and stabilizability properties.

Definition 6.2. *The semigroup $(T(t))_{t\geq 0}$ is said to be* exponentially stable *if there exist positive constants M and α such that*

$$||T(t)|| \leq M\mathrm{e}^{-\alpha t}, \quad t \geq 0. \tag{6.12}$$

We say that $(T(t))_{t\geq 0}$ is β-exponentially stable *if (6.12) holds for some $\alpha > -\beta$.*

If there exists an $F \in L(H;U)$ such that $A + BF$ generates an exponentially stable semigroup $(T_{BF}(t))_{t\geq 0}$, then the pair (A, B) is said to be exponentially stabilizable. *If $(T_{BF}(t))_{t\geq 0}$ is β-exponentially stable then we say that the pair (A, B) is* β-exponentially stabilizable.

Furthermore we will use the following spectral property.

Definition 6.3. *The infinitesemal generator A satisfies the* spectrum decomposition assumption *at $\delta \in \mathbb{R}$ if $\sigma_\delta^+(A)$ is bounded and separated from $\sigma_\delta^-(A)$ in such a way that a rectifiable, simple, closed curve Γ_δ can be drawn so as to enclose an open set containing $\sigma_\delta^+(A)$ in its interior and $\sigma_\delta^-(A)$ in its exterior. Here*

$$\begin{aligned}\sigma_\delta^+(A) &:= \sigma(A) \cap \overline{\mathbb{C}_\delta^+}; \quad \mathbb{C}_\delta^+ := \{\lambda \in \mathbb{C} : \mathrm{Re}\,(\lambda) > \delta\}\,,\\ \sigma_\delta^-(A) &:= \sigma(A) \cap \mathbb{C}_\delta^-; \quad \mathbb{C}_\delta^- := \{\lambda \in \mathbb{C} : \mathrm{Re}\,(\lambda) < \delta\}\,.\end{aligned}$$

If A satisfies the spectrum decomposition assumption at δ then we may define the spectral projection P_δ by

$$P_\delta z := \frac{1}{2\pi \mathrm{i}} \int_{\Gamma_\delta} R(\lambda, A) z \, d\lambda,$$

where Γ_δ is traversed once in counterclockwise direction. This induces the decomposition

$$H = H_\delta^+ \oplus H_\delta^-, \quad \text{where } H_\delta^+ := P_\delta H \text{ and } H_\delta^- := (I - P_\delta)H.$$

We may decompose A, $(T(t))_{t\geq 0}$ and a bounded linear operator $B \in L(U; H)$ as

$$A = \begin{pmatrix} A_\delta^+ & 0 \\ 0 & A_\delta^- \end{pmatrix}, \quad T(t) = \begin{pmatrix} T_\delta^+(t) & 0 \\ 0 & T_\delta^-(t) \end{pmatrix}, \quad B = \begin{pmatrix} B_\delta^+ \\ B_\delta^- \end{pmatrix},$$

with $B_\delta^+ := P_\delta B \in L(U; H_\delta^+)$ and $B_\delta^- := (I - P_\delta)B \in L(U; H_\delta^-)$.

In order to state the main result of this section, we will need the notion of controllability of finite dimensional systems. Let $C \in \mathbb{R}^{n \times n}$ and $D \in \mathbb{R}^{n \times k}$. Recall that the pair (C, D) is called *controllable* if

$$\mathrm{rank}\left[D, CD, \ldots, C^{n-1}D\right] = n.$$

Here $[T_1, T_2, \ldots, T_n]$ denotes the concatenation of all the columns of the matrices $T_1, \ldots, T_n$.

We recall the following theorem (see [20], Theorem 5.2.6).

Theorem 6.4. *Let A be the infinitesemal generator of a strongly continuous semigroup $(T(t))_{t\geq 0}$ on a Hilbert space H and let $B \in L(U; H)$ have finite rank, with U a Hilbert space. Then the following are equivalent:*

(i) The pair (A, B) is β-exponentially stabilizable;

(ii) A satisfies the spectrum decomposition assumption at β, H_β^+ is finite dimensional, $(T_\beta^-(t))_{t\geq 0}$ is β-exponentially stable and the finite dimensional pair (A_β^+, B_β^+) is controllable.

Note that if A generates an eventually compact semigroup then A satisfies the spectrum decomposition assumption with H_β^+ finite dimensional for any $\beta \in \mathbb{R}$ (see Section 5.5.1).

We can now state the main result of this section.

Theorem 6.5. *Suppose A satisfies the spectrum decomposition assumption at $\delta = 0$ with $\dim H_0^+ < \infty$, and that $B \in L(U;H)$ has finite rank. If the pair (A, BF) is stochastically stable for some $F \in L(H;U)$, then (A, B) is exponentially stabilizable.*

Corollary 6.6. *Suppose A satisfies the spectrum decomposition assumption at $\delta = 0$ with $\dim H_0^+ < \infty$, and that $B \in L(U;H)$ has finite rank. Furthermore suppose that the solution of*

$$dX(t) = AX(t)\ dt + BX(t)\ dM(t)$$

is asymptotically stable. Then the system (A, B) is exponentially stabilizable.

Proof (of corollary): By taking $F = I$, we see that (A, B) is stochastically stabilizable. Now apply Theorem 6.5. □

In order to prove Theorem 6.5, we need some further notions and results from systems theory. See [80] for details.

Let, for the discussion below, $C \in \mathbb{R}^{n\times n}$ and $D \in \mathbb{R}^{n\times k}$.

$\lambda \in \mathbb{C}$ is called (C, D)-*controllable* if

$$\operatorname{rank}[C - \lambda I, D] = n,$$

Note that if $\lambda \notin \sigma(C)$, then λ is always (C, D)-controllable.

The system $(\overline{C}, \overline{D})$ is said to be *isomorphic* to (C, D) if there exists an invertible matrix S such that

$$\overline{C} = S^{-1}CS, \quad \overline{D} = S^{-1}D.$$

Lemma 6.7. *Suppose (C, D) is not controllable and $D \neq 0$. Then there exist $(\overline{C}, \overline{D})$ isomorphic to (C, D) such that*

$$\overline{C} = \begin{bmatrix} C_{11} & C_{12} \\ 0 & C_{22} \end{bmatrix}, \quad \overline{D} = \begin{bmatrix} D_1 \\ 0 \end{bmatrix},$$

and such that (C_{11}, D_1) is controllable.

Furthermore $\lambda \in \mathbb{C}$ is (C, D)-controllable if and only if $\lambda \notin \sigma(C_{22})$.

Proof (of Theorem 6.5): Let $(X(t))_{t\geq 0}$ denote the unique solution to

$$X(t) = T(t)x + \int_0^t T(t-s)BFX(s)\ dM(s), \tag{6.13}$$

with $x \in H$.

Let $P := P_0$, $H^+ := H_0^+$, $H^- := H_0^-$ etc. be as defined above and let $n := \dim H^+ < \infty$. Choose an arbitrary basis of H^+ in order to identify H^+ and $\mathbb{R}^n$. Note that

$$dPX(t) = A^+X(t)\ dt + B^+FX(t)\ dM(t).$$

Suppose that A^+ has an eigenvalue $\lambda \in \mathbb{C}$ with Re $(\lambda) \geq 0$ which is *uncontrollable*, so

$$\operatorname{rank}\begin{pmatrix} A^+ - \lambda I & B^+ \end{pmatrix} < n.$$

By Lemma 6.7 there exists a system $(\overline{A}, \overline{B})$ isomorphic to (A^+, B^+) such that

$$\overline{A} = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}, \quad \overline{B} = \begin{pmatrix} B_1 \\ 0 \end{pmatrix},$$

and, for some $S \in GL(\mathbb{R}^n)$,

$$S^{-1}A^+S = \overline{A}, \quad S^{-1}B^+ = \overline{B}.$$

with $\lambda \in \sigma(A_{22})$. Let x be a generalized eigenvector of A^+ corresponding to eigenvalue λ and let $(X(t))_{t\geq 0}$ be the solution of (6.13) for this choice of x.

Now PX satisfies

$$\begin{aligned} dS^{-1}PX(t) &= S^{-1}A^+PX(t)\ dt + S^{-1}B^+FX(t)\ dM(t) \\ &= \overline{A}S^{-1}PX(t)\ dt + \overline{B}FX(t)\ dM(t). \end{aligned}$$

Let Q denote projection on the generalized eigenspace E_λ of A_{22} corresponding to eigenvalue λ, and note that Q commutes with A_{22} and $Q\overline{B} = 0$. Hence

$$dQS^{-1}PX(t) = A_{22}QS^{-1}PX(t)\ dt, \quad QS^{-1}PX(0) = QS^{-1}x,$$

so

$$QS^{-1}PX(t) = \exp(A_{22}t)QS^{-1}x.$$

Since $\sigma(\overline{A}|_{E_\lambda}) = \{\lambda\}$, we have

$$||X(t)|| \geq k||QS^{-1}PX(t)|| \geq \tilde{k}e^{\lambda t}|S^{-1}x|,$$

for some positive constants $k, \tilde{k}$. This shows that $(X(t))_{t\geq 0}$ is not stochastically stable if (A^+, B^+) has an uncontrollable eigenvalue λ with Re $\lambda \geq 0$, or equivalently $(X(t))_{t\geq 0}$ is not stochastically stable if (A^+, B^+) is not controllable. But, by Theorem 6.4, controllability of (A^+, B^+) (combined with the assumptions in the formulation of the theorem) is equivalent to (A, B) being exponentially stabilizable. □

6.4 Tools

In the previous section we discussed a necessary condition for a linear stochastic evolution with multiplicative noise to be exponentially stable. In this section we will introduce some tools which we will use to establish sufficient conditions for this property.

6.4.1 Law of large numbers for martingales

Below we will prove a special case of the *law of large numbers for martingales*. This is not a new result; a more general formulation can be found for example in [56], Theorem 3.4 but is stated there with neither proof nor reference. It seems to be a folklore result: a proof was nowhere to be found but the result is used now and then. We provide a proof of our own here. A proof for the discrete time case may be found in [79].

First we need the exponential martingale inequality:

Proposition 6.8 (Exponential martingale inequality). *Let $(M(t))_{t\geq 0}$ be a continuous local martingale with $M(0) = 0$. Then for all $\varepsilon, \delta > 0$,*

$$\mathbb{P}\left(\sup_{t\geq 0} M(t) \geq \varepsilon \text{ and } [M](\infty) \leq \delta\right) \leq \mathrm{e}^{-\frac{1}{2}\varepsilon^2/\delta}.$$

Proof: Fix $\varepsilon > 0$ and set $T := \inf\{t \geq 0 : M(t) \geq \varepsilon\}$. Fix $\theta \geq 0$ and set

$$Z(t) := \exp\left\{\theta M(t)^T - \tfrac{1}{2}\theta^2 [M](t)^T\right\}.$$

Then Z is a continuous local martingale, and

$$|Z| \leq \mathrm{e}^{\theta\varepsilon}, \quad \text{a.s.}$$

so Z is a bounded martingale, and by optional stopping

$$\mathbb{E}\left[Z(\infty)\right] = \mathbb{E}\left[Z(0)\right] = 1.$$

For $\delta > 0$,

$$\mathbb{P}\left(\sup_{t\geq 0} M(t) \geq \varepsilon \text{ and } [M](\infty) \leq \delta\right) \leq \mathbb{P}\left(Z(\infty) \geq \mathrm{e}^{\theta\varepsilon - \frac{1}{2}\theta^2\delta}\right) \leq \mathrm{e}^{-\theta\varepsilon + \frac{1}{2}\theta^2\delta}.$$

Now optimize over θ. □

Remark 6.9. See also Theorem 23.17 in [42] for exponential inequalities for local martingales with bounded jumps.

Corollary 6.10. *Let $(M(t))_{t\geq 0}$ be a continuous local martingale with $M(0) = 0$. Then for all $\varepsilon, \delta > 0$,*

$$\mathbb{P}\left(\sup_{t\geq 0} |M(t)| \geq \varepsilon \text{ and } [M](\infty) \leq \delta\right) \leq 2\mathrm{e}^{-\frac{1}{2}\varepsilon^2/\delta}.$$

Theorem 6.11 (Law of large numbers for martingales). *Let $(M(t))_{t\geq 0}$ be a continuous local martingale in $\mathbb{R}$ with $M(0) = 0$. If*

$$\limsup_{t\to\infty} \frac{[M](t)}{t} < \infty, \quad a.s., \tag{6.14}$$

then

$$\lim_{t\to\infty} \frac{M(t)}{t} = 0, \quad a.s.$$

PROOF: Let $k, m \in \mathbb{N}$. For $n \in \mathbb{N}$, let E_n denote the event

$$E_n := \left\{ \sup_{t\geq 0} |M^{n+1}(t)| \geq \frac{n}{m} \text{ and } [M^{n+1}](\infty) \leq 2kn \right\}.$$

Here M^n is the continuous local martingale obtained by stopping M at time n.

Then by the exponential martingale inequality,

$$\mathbb{P}(E_n) \leq 2\mathrm{e}^{-\frac{n}{4km^2}}, \quad n \in \mathbb{N}.$$

Hence by Borel-Cantelli, $\mathbb{P}(E_n^c,$ eventually as $n \to \infty) = 1$. Let

$$\widetilde{\Omega}_{k,m} := (E_n^c, \text{ eventually as } n \to \infty)$$

and

$$\Omega_k := \{[M](t) \leq kt \text{ for all } t \geq 0\} \quad \text{for } k \in \mathbb{N}.$$

On Ω_k, we have that

$$\frac{[M]^{n+1}(\infty)}{n} = \frac{[M](n+1)}{n} \leq \frac{(n+1)k}{n} \leq 2k, \quad n \in \mathbb{N},$$

so on $\Omega_k \cap \widetilde{\Omega}_{k,m}$ we have that

$$\sup_{t\geq 0} |M^{n+1}(t)| < \frac{n}{m}, \quad \text{eventually as } n \to \infty.$$

In particular, on $\Omega_k \cap \widetilde{\Omega}_{k,m}$, for $N \in \mathbb{N}$ large enough and $t \in [n, n+1]$, for $n > N$, $n \in \mathbb{N}$,

$$\frac{|M(t)|}{t} = \frac{|M^{n+1}(t)|}{t} \leq \frac{|M^{n+1}(t)|}{n} < \frac{1}{m},$$

that is

$$\frac{|M(t)|}{t} < \frac{1}{m} \quad \text{for } t \text{ large enough.}$$

Note that $\widetilde{\Omega}_k := \cap_m \widetilde{\Omega}_{k,m}$ has full measure and on $\Omega_k \cap \widetilde{\Omega}_k$, for all $m \in \mathbb{N}$ we have

$$\frac{|M(t)|}{t} < \frac{1}{m} \quad \text{for } t \text{ large enough.}$$

Note that, by (6.14), for all $\gamma > 0$ there exists an $k \in \mathbb{N}$ such that $\mathbb{P}(\Omega_k) \geq 1 - \gamma$. Therefore $\widetilde{\Omega} := \cup_k \widetilde{\Omega}_k$ has full measure and on $\widetilde{\Omega}$, for all $m \in \mathbb{N}$,

$$\frac{|M(t)|}{t} < \frac{1}{m} \quad \text{for } t \text{ large enough.}$$

□

6.4.2 Approximation of solutions of stochastic differential equations

To be able to extend results from the case of uniformly continuous semigroups to the more general case of strongly continuous semigroup, we will need some results on approximation of solutions of SDEs.

Definition 6.12. *Let X be a Banach space. Let A be the infinitesemal generator of a strongly continuous semigroup $(S(t))_{t\geq 0}$ in $L(X)$. Let $(T_n(t))_{t\geq 0, n\in\mathbb{N}}$ be a sequence of strongly continuous semigroups in $L(X)$ with generators $(A_n)_{n\in\mathbb{N}}$.*

Then (A_n) is called an approximation *of A if*

(i) there exists $M > 0$ and $\omega \in \mathbb{R}$ such that

$$||T(t)|| \vee \sup_{n\in\mathbb{N}} ||T_n(t)|| \leq M\mathrm{e}^{\omega t} \quad \text{for all } t \geq 0, \tag{6.15}$$

and

(ii) we have that $T_n(t)x \to T(t)x$ as $n \to \infty$ for all $x \in X$, uniformly in t on compact sets.

An approximation (A_n) of A is called a bounded approximation *if $A_n \in L(X)$ for all $n \in \mathbb{N}$.*

We will not distinguish between an approximation (A_n) of A or an approximation (T_n) of the corresponding semigroup T.

Equivalent conditions for (A_n) to be an approximation of A are given by the Trotter-Kato theorem, see [32], Theorem III.4.8. A sufficient condition for a sequence (A_n) to be an approximation of A is that (6.15) holds, and that $A_n x \to Ax$ for all $x \in D$, where D is a core for A.

Yosida approximation

An important example of bounded approximations is the *Yosida approximation* which we will discuss here.

Let X be a Banach space. In this example, $A : \mathfrak{D}(A) \to X$ is the infinitesemal generator of a strongly continuous semigroup $(S(t))_{t\geq 0}$ on X satisfying

$$||S(t)|| \leq M\mathrm{e}^{\omega t}, \quad t \geq 0,$$

where $M \geq 1$ and $\omega \in \mathbb{R}$.

Recall the notions of the resolvent set of A,

$$\rho(A) := \{\lambda \in \mathbb{C} : \lambda - A \text{ has a bounded inverse}\},$$

and the resolvent of A,

$$R(\lambda, A) := (\lambda - A)^{-1}, \quad \lambda \in \rho(A).$$

Define the *Yosida approximation* of A by

$$A_n := AJ_n = nAR(n, A) = n^2 R(n, A) - nI, \quad n \in \mathbb{N} \cap \rho(A),$$

where $J_n := nR(n, A)$. These $(A_n)_{n\in\mathbb{N}}$ are bounded operators and therefore generate uniformly continuous semigroups which we denote by $(S_n(t))_{t\geq 0}$, $n \in \mathbb{N}$. Furthermore $A_n x \to Ax$ for all $x \in \mathfrak{D}(A)$.

By [24], Theorem A.2,

$$||S_n(t)|| \leq M \mathrm{e}^{\frac{\omega n t}{n-\omega}}, \quad t \geq 0.$$

Note that

$$\frac{\omega n}{n - \omega} \to \omega.$$

In particular, for all $\widetilde{\omega} > \omega$ there exists an $N \in \mathbb{N}$ such that for and all $n > N$, we have

$$||S_n(t)|| \leq M \mathrm{e}^{\widetilde{\omega} t}, \quad t \geq 0. \tag{6.16}$$

If we combine the Trotter-Kato approximation theorem ([32], Theorem III.4.8) with (6.16), we obtain the following lemma:

Lemma 6.13. *$S_n(t)x \to S(t)x$ for all $x \in X$, uniformly for $t \in [0, T]$ with $T > 0$.*

Proposition 6.14. *Suppose S is a strongly continuous semigroup with infinitesemal generator A, M is a continuous cylindrical martingale of stationary covariance with RKHS $\mathcal{H}$, $p > 2$ and $X_0 \in L^p(\Omega, \mathcal{F}_0; H)$. Suppose $(A_n)_{n\in\mathbb{N}}$ are approximations of A. Suppose $F : H \to H$ and $G : H \to L_{\mathrm{HS}}(\mathcal{H}; H)$ are globally Lipschitz. Let X be the unique mild solution to*

$$dX(t) = (AX + F(X))\, dt + G(X)\, dM(t), \quad X(0) = X_0,$$

and X_n the unique mild solution to

$$dX(t) = (A_n X_n + F(X_n))\, dt + G(X_n)\, dM(t), \quad X(0) = X_0.$$

Then for all $T > 0$,

$$\mathbb{E} \sup_{t\in[0,T]} |X(t) - X_n(t)|^p \to 0 \quad \text{as} \quad n \to \infty.$$

Moreover, there exists a sequence $(n_k)_{k\in\mathbb{N}}$ in $\mathbb{N}$ such that

$$\lim_{k\to\infty} \sup_{t\in[0,T]} |X_{n_k}(t) - X(t)| \quad \text{for all } T > 0, \quad \text{almost surely.}$$

PROOF: Let $\varepsilon > 0$.

$$\begin{aligned}
&\mathbb{E}\left[\sup_{t\in[0,T]} |X(t) - X_n(t)|^p\right] \\
&\leq 3^{p-1}\mathbb{E}\left[\sup_{t\in[0,T]} |S(t)X_0 - S_n(t)X_0|^p\right] \\
&\quad + 3^{p-1}\mathbb{E}\left[\sup_{t\in[0,T]} \left|\int_0^t S(t-s)F(X(s)) - S_n(t-s)F(X_n(s))\ ds\right|^p\right] \\
&\quad + 3^{p-1}\mathbb{E}\left[\sup_{t\in[0,T]} \left|\int_0^t S(t-s)G(X(s)) - S_n(t-s)G(X_n(s),s)\ dM(s)\right|^p\right].
\end{aligned}$$

First term. For the first term we have that

$$\lim_{n\to\infty} \sup_{t\in[0,T]} |S(t)X_0 - S_n(t)X_0|^2 \to 0, \quad \text{almost surely.}$$

Since $(S(t))_{t\in[0,T]}$ and $(S_n(t))_{t\in[0,T]}$ are uniformly bounded in operator norm, by dominated convergence,

$$3^{p-1}\mathbb{E}\left[\sup_{t\in[0,T]} |S(t)X_0 - S_n(t)X_0|^2\right] < \varepsilon$$

for n large enough.

Second term. Note that

$$\begin{aligned}
&\sup_{s\in[0,t]} \left|\int_0^s S(s-r)F(X(r)) - S_n(s-r)F(X_n(r))\ dr\right|^p \\
&\leq \sup_{s\in[0,t]} s^{p-1} \int_0^s |S(s-r)F(X(r)) - S_n(s-r)F(X_n(r))|^p\ dr \\
&\leq \sup_{s\in[0,t]} (2s)^{p-1} \int_0^s |S_n(s-r)\left(F(X(r)) - F(X_n(r))\right)|^p\ dr \\
&\qquad + \sup_{s\in[0,t]} (2s)^{p-1} \int_0^s |S_n(s-r)F(X(r)) - S(s-r)F(X(r))|^p\ dr.
\end{aligned}$$

with

$$\begin{aligned}
&\sup_{s\in[0,t]} \int_0^s |S_n(s-r)\left(F(X(r)) - F(X_n(r))\right)|^p\ dr \\
&\leq k_1 \int_0^t \sup_{r\in[0,s]} |X(r) - X_n(r)|^p\ ds
\end{aligned}$$

where the constant $k_1 > 0$ may be chosen such that the inequality holds for all $t \in [0,T]$. Furthermore by dominated convergence, for n large enough,

$$\mathbb{E} \sup_{s\in[0,t]} (2s)^{p-1} \int_0^s |S_n(s-r)F(X(r)) - S(s-r)F(X(r))|^p \, dr < \varepsilon.$$

Third term. For the third term, employing Proposition 2.29 (ii) twice, for n large enough,

$$\begin{aligned}
&\mathbb{E} \sup_{s\in[0,t]} \left| \int_0^s S(s-r)G(X(r)) - S_n(s-r)G(X_n(r)) \, dM(r) \right|^p \\
&\leq 2^{p-1}\mathbb{E} \sup_{s\in[0,t]} \left| \int_0^s S(s-r)G(X(r)) - S_n(s-r)G(X(r)) \, dM(s) \right|^p \\
&\qquad + 2^{p-1}\mathbb{E} \sup_{s\in[0,t]} \left| \int_0^s S_n(s-r)(G(X(r)) - G(X_n(r))) \, dM(s) \right|^p \\
&\leq \varepsilon + k_2 \mathbb{E} \int_0^t \sup_{r\in[0,s]} |X(r) - X_n(r)|^p \, dt,
\end{aligned}$$

where the constant $k_2 > 0$ may be chosen in such a way that the inequality holds for all $t \in [0,T]$.

Combining all the terms, for $t \in [0,T]$ and for n large enough,

$$\begin{aligned}
&\mathbb{E} \sup_{s\in[0,t]} |X(s) - X_n(s)|^p \\
&\leq 3\varepsilon + (k_1 + k_2) \int_0^t \mathbb{E} \sup_{r\in[0,s]} |X(r) - X_n(r)|^p \, ds.
\end{aligned}$$

By Gronwall's lemma therefore, for $t \in [0,T]$,

$$\mathbb{E} \sup_{s\in[0,t]} |X(s) - X_n(s)|^p \leq 3\varepsilon \exp((k_1 + k_2)t),$$

and we may let $\varepsilon \downarrow 0$ to obtain the first claim.

It follows that for every $T > 0$ there exists a subsequence $(n(k))_{k\in\mathbb{N}}$ such that

$$\lim_{k\to\infty} \sup_{t\in[0,T]} |X(t) - X_{n(k)}(t)| = 0, \quad \text{a.s.}$$

Now define $\Omega_1 \subset \Omega$, $\mathbb{P}(\Omega_1) = 1$ and a subsequence $(n_1(k))_{k\in\mathbb{N}}$ such that

$$\lim_{k\to\infty} \sup_{t\in[0,1]} |X_{n_1(k)}(t) - X(t)| = 0 \quad \text{on } \Omega_1.$$

Define recursively, for $m \in \mathbb{N}$, $m \geq 2$, sets $\Omega_m \subset \Omega$, $\mathbb{P}(\Omega_m) = 1$, and further subsequences $(n_m(k))_{k\in\mathbb{N}}$ of $(n_{m-1}(k))_{k\in\mathbb{N}}$ such that

$$\lim_{k\to\infty} \sup_{t\in[0,m]} |X_{n_m(k)}(t) - X(t)| = 0 \quad \text{on } \Omega_m.$$

Let $\tilde{\Omega} := \cap_{m\in\mathbb{N}}\Omega_m$ (so $\mathbb{P}(\tilde{\Omega}) = 1$), and consider the subsequence $(n_k(k))_{k\in\mathbb{N}}$. Let $\omega \in \tilde{\Omega}$, $\varepsilon > 0$ and $M \in \mathbb{N}$. Take $K > 0$ such that

$$\sup_{t\in[0,M]} |X_{n_M}(k)(t) - X(t)|(\omega) < \varepsilon \quad \text{for all } k \geq K.$$

Then for $k \geq K \vee M$, $n_k(k) \geq n_M(k)$, and $n_k(k) \in (n_M(l))_{l\in\mathbb{N}}$, and therefore

$$\sup_{t\in[0,M]} |X_{n_k}(k)(t) - X(t)|(\omega) < \varepsilon.$$

This proves the second claim: There exists a set $\tilde{\Omega}$ with $\mathbb{P}(\tilde{\Omega}) = 1$ and a subsequence $(n_k(k))_{k\in\mathbb{N}}$ of $\mathbb{N}$ such that on $\tilde{\Omega}$, we have that for any $M \in \mathbb{N}$ (and hence any $T > 0$),

$$\lim_{k\to\infty} \sup_{t\in[0,M]} |X_{n_k(k)}(t) - X(t)| = 0.$$

$\square$

6.5 Estimate on the Lyapunov exponent of non-degenerate diffusions

In this section we will establish a pathwise stability result for the case where the noise occurs in the stochastic differential equation in a non-degenerate way.

Recall that a linear operator $A : \mathfrak{D}(A) \to H$ on H is called *dissipative* when

$$\langle Ax, x\rangle \leq 0 \quad \text{for all } x \in \mathfrak{D}(A).$$

The generator A of a strongly continuous semigroup $(S(t))_{t\geq 0}$ is dissipative if and only if $(S(t))_{t\geq 0}$ is a contraction semigroup.

Theorem 6.15. *Suppose $A : \mathfrak{D}(A) \to H$ is the generator of a strongly continuous semigroup $(S(t))_{t\geq 0}$, and suppose $F, G : H \times \mathbb{R}_+ \to H$ are Lipschitz continuous with $F(0,t) = G(0,t) = 0$, $t \geq 0$. Let $(W(t))_{t\geq 0}$ be a one-dimensional standard Brownian motion. Let $X(t;x)$ denote the unique solution $x : \Omega \times [0,\infty) \to \mathbb{R}$ to the stochastic differential equation*

$$X(t) = S(t)x + \int_0^t S(t-s)F(X(s),s)\, ds + \int_0^t S(t-s)G(X(s),s)\, dW(s). \quad (6.17)$$

Suppose there exists a $k > 0$ such that $A - kI$ is dissipative. Suppose finally that there exists $\gamma \in \mathbb{R}$ such that

$$\frac{|G(x,t)|^2}{|x|^2} - \frac{2\langle G(x,t), x\rangle^2}{|x|^4} \leq \gamma \quad \textit{for all } x \in H, x \neq 0 \textit{ and all } t \geq 0. \quad (6.18)$$

Then

$$\limsup_{t\to\infty} \frac{1}{t} \log |X(t;x)|^2 \le 2(k + [F]_{\mathrm{Lip}}) + \gamma.$$

In particular, if $2(k + [F]_{\mathrm{Lip}}) + \gamma < 0$ then the solution of (6.17) is almost surely exponentially stable.

Remark 6.16. Note that, for a stabilizing effect of the noise, we require that $\gamma < 0$ in (6.18). However, if there exists an $X \in H, x \neq 0$ such that $G(X,t) = 0,\ t \ge 0$, then we have that $\gamma \ge 0$. This shows that we can only hope to establish a stabilizing effect from Theorem 6.15 by using a nondegenerate noise term.

Proof (of Theorem 6.15): Let $\tilde{A} := A - kI$ be the generator of the contraction semigroup $(\tilde{S}(t))_{t\ge 0}$, and let $\tilde{F}(x,t) := F(x,t) + kx$ and rewrite (6.17) into

$$X(t) = \tilde{S}(t)x + \int_0^t \tilde{S}(t-s)\tilde{F}(X(s),s)\ ds + \int_0^t \tilde{S}(t-s)G(X(s),s)\ dW(s).$$

Note that $\tilde{F}$ is Lipschitz continuous with $[\tilde{F}]_{\mathrm{Lip}} = [f]_{\mathrm{Lip}} + k$.

Define the operators $(A_n)_{n\in\mathbb{N}}$ to be the Yosida approximations of $\tilde{A}$, and let $(X_n)_{n\in\mathbb{N}}$ be the solutions to

$$X_n(t) = S_n(t)x + \int_0^t S_n(t-s)\tilde{F}(X_n(s),s)\ ds + \int_0^t S_n(t-s)G(X_n(s),s)\ dW(s).$$

Fix $n \in \mathbb{N}$. Define the real valued stochastic process

$$Y_n(t) := |X_n(t)|^2.$$

Using Itô's formula,

$$\begin{aligned} dY_n(t) &= 2\langle X_n(t), A_nX_n(t) + \tilde{F}(X_n(t),t)\rangle\ dt \\ &\quad + 2\langle X_n(t), G(X_n(t),t)\rangle\ dW(t) + |G(X_n(t),t)|^2\ dt. \end{aligned}$$

Define

$$\delta(t) := \begin{cases} \frac{|G(X(t),t)|^2}{|X(t)|^2} & \text{for } X(t) \neq 0 \\ \delta_0 & \text{for } X(t) = 0 \end{cases} \quad \text{and}$$

$$\varepsilon(t) := \begin{cases} \frac{2\langle X(t), G(X(t),t)\rangle}{|X(t)|^2} & \text{for } X(t) \neq 0 \\ \varepsilon_0 & \text{for } X(t) = 0, \end{cases}$$

for some δ_0, ε_0 such that

$$\delta_0 - \tfrac{1}{2}\varepsilon_0^2 \le \gamma,$$

and for $n \in \mathbb{N}$ define

$$\delta_n(t) := \begin{cases} \frac{|G(X_n(t),t)|^2}{|X_n(t)|^2} & \text{for } X_n(t) \neq 0 \\ \delta(t) & \text{for } X_n(t) = 0, \end{cases} \quad \text{and}$$

$$\varepsilon_n(t) := \begin{cases} \frac{2\langle X_n(t), G(X_n(t),t)\rangle}{|X_n(t)|^2} & \text{for } X_n(t) \neq 0 \\ \varepsilon(t) & \text{for } X_n(t) = 0. \end{cases}$$

Then

$$dY_n(t) = 2\langle X_n(t), A_n X_n(t) + \tilde{F}(X_n(t),t)\rangle\ dt + \varepsilon_n(t) Y(t)\ dW(t) + \delta_n(t) Y(t)\ dt.$$

Note, by the Lipschitz continuity of G, that δ_n and ε_n are uniformly bounded in $n \in \mathbb{N}$ and $t \geq 0$.

Using variation of constants,

$$Y_n(t) = \Phi_n(t) \left(Y(0) + 2\int_0^t \Phi_n^{-1}(s) \langle X_n(s), A_n X_n(s) + \tilde{F}(X_n(s), s)\rangle\ ds \right),$$

where

$$\Phi_n(t) = \exp\left(\int_0^t \delta_n(s) - \tfrac{1}{2}\varepsilon_n(s)^2\ ds + \int_0^t \varepsilon_n(s)\ dW(s) \right).$$

By the Lipschitz continuity of $\tilde{F}$ and dissipativity of $\tilde{A}$,

$$Y_n(t) \leq \Phi_n(t) \left(Y(0) + 2\alpha \int_0^t \Phi_n^{-1}(s) Y_n(s)\ ds \right),$$

where $\alpha = [\tilde{F}]_{\mathrm{Lip}} = [F]_{\mathrm{Lip}} + k$. Write $Z_n(t) := \Phi_n^{-1}(t) Y_n(t)$, $t \geq 0$, $n \in \mathbb{N}$, then

$$Z_n(t) \leq Y(0) + 2\alpha \int_0^t Z_n(s)\ ds$$

so that by Gronwall's lemma,

$$Z_n(t) \leq y(0) \exp(2\alpha t) \quad \text{almost surely, for } t \geq 0.$$

Hence

$$Y_n(t) = \Phi_n(t) Z_n(t) \leq \Phi_n(t) Y(0) \exp(2\alpha t) \quad \text{almost surely for } t \geq 0.$$

We find that, almost surely, for $t \geq 0$, $n \in \mathbb{N}$,

$$\frac{1}{t} \log |X_n(t)|^2 = \frac{1}{t} \log Y_n(t) \leq \frac{1}{t} \log Y(0) + \frac{1}{t} \log \Phi_n(t) + 2\alpha. \tag{6.19}$$

For $n \in \mathbb{N}$,

$$\frac{1}{t}\log \Phi_n(t) = \frac{1}{t}\left(\int_0^t \delta_n(s) - \tfrac{1}{2}\varepsilon_n^2(s)\; ds + \int_0^t \varepsilon_n(s)\; dW(s)\right)$$

Let, for $t \geq 0$,

$$\widetilde{\Phi}_n(t) := \exp\left(\int_0^t \left\{\delta_n(s) - \tfrac{1}{2}\varepsilon_n(s)^2\right\} \mathbb{1}_{\{X(s)\neq 0\}}\; ds + \int_0^t \varepsilon_n(s)\mathbb{1}_{\{X(s)\neq 0\}}\; dW(s)\right)$$

and

$$\widetilde{\Phi}(t) := \exp\left(\int_0^t \left\{\delta(s) - \tfrac{1}{2}\varepsilon(s)^2\right\} \mathbb{1}_{\{X(s)\neq 0\}}\; ds + \int_0^t \varepsilon(s)\mathbb{1}_{\{X(s)\neq 0\}}\; dW(s)\right).$$

Suppose $t > 0$. By Lemma 6.14, there exists a sequence $(n_k)_{k\in\mathbb{N}}$ in $\mathbb{N}$ such that

$$\sup_{s\in[0,t]} |X_{n_k}(s) - X(s)| \to 0 \quad \text{as } k\to\infty, \quad \text{almost surely.} \tag{6.20}$$

Hence for all $s \in [0,t]$,

$$|\delta_{n_k}(s) - \delta(s)|\mathbb{1}_{\{x(s)\neq 0\}} \vee |\varepsilon_{n_k}(s) - \varepsilon(s)|\mathbb{1}_{\{x(s)\neq 0\}} \to 0 \quad \text{as } k\to\infty, \quad \text{almost surely,}$$

and since $(\delta_n), \delta, (\varepsilon_n)$, and ε are uniformly bounded, we see that, by dominated convergence,

$$\begin{aligned}
&\mathbb{E}\left[\left|\int_0^t (\varepsilon_{n_k}(s) - \varepsilon(s))\mathbb{1}_{\{X(s)\neq 0\}}\; dW(s)\right|^2\right] \\
&= \mathbb{E}\left[\int_0^t |\varepsilon_{n_k}(s) - \varepsilon(s)|^2 \mathbb{1}_{\{X(s)\neq 0\}}\; ds\right] \to 0 \quad \text{as } k\to\infty.
\end{aligned}$$

In a similar way

$$\mathbb{E}\left[\left|\int_0^t \left(\left\{\delta_{n_k}(s) - \tfrac{1}{2}\varepsilon_{n_k}^2(s)\right\} - \left\{\delta(s) - \tfrac{1}{2}\varepsilon^2(s)\right\}\right)\mathbb{1}_{\{X(s)\neq 0\}}\; ds\right|^2\right] \to 0 \quad \text{as } k\to\infty.$$

Therefore, there exists a subsequence $(n_{k_l})_{l\in\mathbb{N}}$ such that

$$\left|\widetilde{\Phi}_{n_{k_l}}(t) - \widetilde{\Phi}(t)\right| \to 0 \quad \text{as } k\to\infty, \quad \text{almost surely,} \tag{6.21}$$

Note that

$$\{\omega\in\Omega : X(t)\neq 0\} \subset \{\omega\in\Omega : X(s)\neq 0 \text{ for } 0\leq s\leq t\},$$

so from (6.19), (6.20), and (6.21) we may conclude that, since $t > 0$ is arbitrary,

$$\frac{1}{t}\log|X(t)|^2 \leq \frac{1}{t}y(0) + \frac{1}{t}\log\widetilde{\Phi}(t) + 2\alpha, \quad \text{a.s. on } \{X(t)\neq 0\}, \text{ for all } t > 0.$$

Furthermore,

$$\frac{1}{t}\log|X(t)|^2 = -\infty \quad \text{on } \{X(t)=0\}, \text{ for all } t>0.$$

It follows that

$$\frac{1}{t}\log|X(t)|^2 \le \frac{1}{t}Y(0) + \frac{1}{t}\log\Phi(t) + 2\alpha, \quad \text{almost surely, for all } t>0,$$

where

$$\Phi(t) = \exp\left(\int_0^t \delta(s) - \tfrac{1}{2}\varepsilon(s)^2 \; ds + \int_0^t \varepsilon(s) \; dW(s)\right).$$

By (almost sure) continuity of the paths of X and Φ, it follows that

$$\frac{1}{t}\log|X(t)|^2 \le \frac{1}{t}Y(0) + \frac{1}{t}\log\Phi(t) + 2\alpha, \quad \text{for all } t>0, \text{almost surely.}$$

Note that

$$\frac{1}{t}\log\Phi(t) = \frac{1}{t}\left(\int_0^t \delta(s) - \tfrac{1}{2}\varepsilon^2(s) \; ds + \int_0^t \varepsilon(s) \; dW(s)\right).$$

By uniform boundedness of ε and the strong law of large numbers for martingales (Theorem 6.11),

$$\frac{1}{t}\int_0^t \varepsilon(s) \; dW(s) \to 0 \quad \text{as } t\to\infty, \quad \text{almost surely.}$$

Furthermore, by assumption (6.18),

$$\delta(t) - \tfrac{1}{2}\varepsilon^2(t) = \frac{|G(X(t),t)|^2}{|X(t)|^2} - \frac{2\langle X(t), G(x(t),t)\rangle^2}{|X(t)|^4} \le \gamma,$$

for (almost all) those (t,ω) for which $X(t)\neq 0$. For (t,ω) such that $X(t)=0$,

$$\delta(t) - \tfrac{1}{2}\varepsilon^2(t) = \delta_0 - \tfrac{1}{2}\varepsilon_0^2 \le \gamma.$$

So

$$\limsup_{t\to\infty}\frac{1}{t}\log\Phi(t) \le \gamma, \quad \text{almost surely,}$$

and the claimed result follows. □

6.5.1 Example: stochastic differential equation in $\mathbb{R}^2$

Consider the following stochastic differential equation in $\mathbb{R}^2$:

$$\begin{cases} dx(t) = Ax(t)\ dt + Bx(t)\ dW(t) \\ x(0) = x_0, \end{cases}$$

with for some $\phi \in [0, 2\pi)$ and $b > 0$,

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad B = bR(\phi) = b \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}.$$

Condition (6.18) gives as sufficient condition for stability

$$\begin{aligned}
& \sup_{x \in \mathbb{R}^2} \frac{|Bx|^2}{|x|^2} - \frac{2\langle Bx, x\rangle^2}{|x|^4} < -2 \\
\iff \quad & \sup_{x \in \mathbb{R}^2} b^2 \left(\frac{|R(\phi)x|^2}{|x|^2} - \frac{2\langle R(\phi)x, x\rangle^2}{|x|^4} \right) < -2 \\
\iff \quad & \sup_{\substack{x \in \mathbb{R}^2 \\ |x|=1}} b^2 \left(\frac{|R(\phi)x|^2}{|x|^2} - \frac{2\langle R(\phi)x, x\rangle^2}{|x|^4} \right) < -2 \\
\iff \quad & \sup_{\substack{x \in \mathbb{R}^2 \\ |x|=1}} b^2 \left(1 - 2\langle R(\phi)x, x\rangle^2\right) < -2 \\
\iff \quad & b^2 \left(1 - 2\cos^2\phi\right) < -2 \\
\iff \quad & b^2 \cos 2\phi > 2.
\end{aligned} \tag{6.22}$$

Since A and B commute in this example, we can also calculate the solution explicitly,

$$x(t) = \exp\left((A - \tfrac{1}{2}B^2)t + BW(t)\right) x_0.$$

Since the eigenvalues of the operator B are $b\exp(\pm \mathrm{i}\phi)$, we find for the maximal Lyapunov exponent

$$\begin{aligned}
\sup_{x_0 \in S^1} \limsup_{t\to\infty} \frac{1}{t} \log |x(t)|^2 &= \sup_{x_0 \in S^1} \limsup_{t\to\infty} \frac{1}{t} \log \left|\exp\left((A - \tfrac{1}{2}B^2)t + BW(t)\right) x_0\right|^2 \\
&= \max_{\lambda \in \sigma(B)} \limsup_{t\to\infty} \frac{1}{t} \log \left|\exp\left((1 - \tfrac{1}{2}\lambda^2)t + \lambda W(t)\right)\right|^2 \\
&= \max_{\lambda \in \sigma(B)} \lim_{t\to\infty} \frac{2}{t} \left(\mathrm{Re}\ (1 - \tfrac{1}{2}\lambda^2)t + \mathrm{Re}\ \lambda W(t)\right) \\
&= 2\max\left\{1 - \tfrac{1}{2}b^2 \mathrm{Re}\ (\exp(2\mathrm{i}\phi)), 1 - \tfrac{1}{2}b^2 \mathrm{Re}\ (\exp(-2\mathrm{i}\phi))\right\} \\
&= 2 - b^2 \cos 2\phi.
\end{aligned}$$

We may conclude that for this example, condition (6.18), or equivalently (6.22) is not only sufficient but also necessary in order to obtain stability. In particular, it is necessary that $\cos(2\phi) > 0$, or equivalently $\phi \in \left(\frac{-\pi}{4}, \frac{\pi}{4}\right) \cup \left(\frac{3\pi}{4}, \frac{5\pi}{4}\right)$.

6.6 Estimate on the Lyapunov exponent of degenerate diffusions

Let H be a real Hilbert space, and consider the linear SDE in mild form

$$X(t) = S(t)x + \sum_{i=1}^{k} \int_0^t S(t-s)B_i X(s)\, dW_i(s), \tag{6.23}$$

where W_i, $i = 1, \ldots, k$, are independent standard Brownian motions in $\mathbb{R}$, $(S(t))_{t\geq 0}$ is a strongly continuous semigroup on H with generator $A : \mathfrak{D}(A) \to H$, $B_i \in L(H)$, $i = 1, \ldots, k$ and $x \in H$.

6.6.1 Bounded case

First we consider the case where $A \in L(H)$, so that $(S(t))_{t\geq 0}$ is a uniformly continuous semigroup. In this case we may apply Itô's formula.

We need the notion of a coercive operator.

Definition 6.17. *An operator $T \in L(H)$ is called* coercive *if $\langle Tx, x\rangle \geq \gamma |x|^2$ for all $x \in H$ and some $\gamma > 0$.*

Lemma 6.18. *Suppose there exists a self-adjoint coercive operator $Q \in L(H)$ and $\lambda \in \mathbb{R}$ such that, for all $x \in H$,*

$$\langle Qx, x\rangle \left[2\langle QAx, x\rangle + \sum_{i=1}^{k} \langle QB_i x, B_i x\rangle - 2\lambda \langle Qx, x\rangle \right] \leq 2 \sum_{i=1}^{k} \langle Qx, B_i x\rangle^2.$$

Let $Y(t) := \langle QX(t), X(t)\rangle$, with X the solution of (6.23).

Then

$$\limsup_{t\to\infty} \frac{1}{t} \log Y(t) \leq 2\lambda, \quad a.s.$$

PROOF: If $x = 0$, then $\mathbb{P}(X(t) = 0) = 1$, $t \geq 0$, and the required estimate holds trivially.

Suppose $x \neq 0$. Then by uniqueness of the solution of SDEs and positiveness of Q, $\mathbb{P}(Y(t) = 0) = 0$ for all $t \geq 0$.

By Itô's formula,

$$\begin{aligned} d\log Y(t) &= \left\{ \frac{1}{Y(t)} \left[2\langle QAX(t), X(t)\rangle + \sum_{i=1}^{k} \langle QB_iX(t), B_iX(t)\rangle \right] \right. \\ &\quad \left. - \frac{2}{Y(t)^2} \sum_{i=1}^{k} \langle QX(t), B_iX(t)\rangle^2 \right\} dt \\ &\quad + \frac{2}{Y(t)} \sum_{i=1}^{k} \langle Qx(t), B_iX(t)\rangle \; dW_i(t) \\ &\leq 2\lambda \; dt + \frac{2}{Y(t)} \sum_{i=1}^{k} \langle QX(t), B_iX(t)\rangle \; dW_i(t). \end{aligned}$$

Now by boundedness of $\frac{\langle QX(t), B_iX(t)\rangle}{\langle QX(t), X(t)\rangle}$ for $i = 1, \ldots, k$ and the law of large numbers for martingales (Theorem 6.11)

$$\frac{1}{t} \int_0^t \frac{2\langle QX(s), B_iX(s)\rangle}{Y(s)} \; dW_i(s) \to 0 \quad (t \to \infty), \quad \text{a.s.}, \quad i = 1, \ldots, k,$$

so

$$\limsup_{t\to\infty} \frac{1}{t} \log Y(t) \leq 2\lambda \quad \text{a.s.}$$

□

Proposition 6.19. *Suppose there exists a self-adjoint matrix $Q \in L(H)$, $\lambda \in \mathbb{R}$ and $b_i \in \mathbb{R}$, $i = 1, \ldots, k$ such that*

$$\left(A + \sum_{i=1}^{k} b_i B_i \right)^* Q + Q \left(A + \sum_{i=1}^{k} b_i B_i \right) + \sum_{i=1}^{k} B_i^* Q B_i + \left(\tfrac{1}{2} \sum_{i=1}^{k} b_i^2 - 2\lambda \right) Q \leq 0. \tag{6.24}$$

Then, with Y as in Lemma 6.18,

$$\limsup_{t\to\infty} \frac{1}{t} \log Y(t) \leq 2\lambda, \quad \text{a.s.}$$

Proof: Note that (by the *abc*-formula), for $i = 1, \ldots, k$,

$$\frac{\langle Qx, B_ix\rangle^2}{\langle Qx, x\rangle^2} + b_i \frac{\langle Qx, B_ix\rangle}{\langle Qx, x\rangle} + \frac{1}{4} b_i^2 \geq 0, \quad \text{for all } x \in H. \tag{6.25}$$

So, by Lemma 6.18, if

$$\begin{aligned} &2\langle QAx, x\rangle + \sum_{i=1}^{k} \langle QB_ix, B_ix\rangle - 2\lambda\langle Qx, x\rangle \\ &\leq -2 \sum_{i=1}^{k} b_i \langle Qx, B_ix\rangle - \frac{1}{2} \sum_{i=1}^{k} b_i^2 \langle Qx, x\rangle \quad \text{for all } x \in H, \end{aligned}$$

then the claimed result holds.

But this is equivalent to the stated condition. □

The next theorem gives a sufficient condition in order for a solution to (6.24) to exist.

Theorem 6.20. *Suppose $D_j \in L(H)$, $j = 1, \ldots, k$, L is the generator of a strongly continuous semigroup $(T(t))_{t\geq 0}$ acting on H such that*

$$||T(t)|| \leq me^{\omega t} \quad \textit{for all } t \geq 0,$$

with $m \geq 1$, $\omega \in \mathbb{R}$, and

$$m^2 \sum_{j=1}^{k} ||D_j||^2 + 2\omega < 0. \tag{6.26}$$

Then for any $M \in L(H)$ there exists a unique solution $Q \in L(H)$ to the equation

$$L^*Q + QL + \sum_{j=1}^{k} D_j^* Q D_j = M, \tag{6.27}$$

which should be interpreted as

$$\langle Qx, Ly\rangle + \langle QLx, y\rangle + \sum_{j=1}^{k} \langle QD_j x, D_j y\rangle = \langle Mx, y\rangle \quad \textit{for all } x, y \in \mathfrak{D}(L). \tag{6.28}$$

This Q also satisfies

$$Q = \int_0^\infty T^*(t) \left(\sum_{j=1}^{k} D_j^* Q D_j - M \right) T(t)\ dt. \tag{6.29}$$

We can estimate the norm of Q by

$$||Q|| \leq -\frac{||M||m^2}{2\omega + m^2 \sum_{j=1}^{k} ||D_j||^2}. \tag{6.30}$$

Furthermore,

(i) if $M = 0$ then $Q = 0$,

(ii) if $M \leq 0$ then $Q \geq 0$, and

(iii) if $M < 0$ then $Q > 0$.

PROOF:

Define a recursion by

$$Q_0 := 0, \quad Q_{i+1} := \int_0^\infty T^*(t) \left(\sum_{j=1}^k D_j^* Q_i D_j - M \right) T(t)\ dt.$$

The recursion is actually a contraction, since

$$\begin{aligned} ||Q_{i+1} - Q_i|| &= \left|\left| \sum_{j=1}^k \int_0^\infty T^*(t) D_j^* (Q_i - Q_{i-1}) D_j T(t)\ dt \right|\right| \\ &\leq m^2 \sum_{j=1}^k ||D_j||^2 \int_0^\infty \mathrm{e}^{2\omega t}\ dt\ ||Q_i - Q_{i-1}|| \\ &= -\frac{m^2 \sum_{j=1}^k ||D_j||^2}{2\omega} ||Q_i - Q_{i-1}||. \end{aligned}$$

Note that the recursion is defined such that Q_{i+1} satisfies

$$L^* Q_{i+1} + Q_{i+1} L = M - \sum_{j=1}^k D_j^* Q_i D_j,$$

a basic result from Lyapunov theory (see [20], Theorem 4.1.23).

Hence there exists a unique fixed point $Q \in L(H)$ that satisfies both (6.27) and

$$Q = \int_0^\infty T^*(t) \left(\sum_{j=1}^k D_j^* Q D_j - M \right) T(t)\ dt.$$

Hence

$$\begin{aligned} ||Q|| &\leq \int_0^\infty m^2 \mathrm{e}^{2\omega t} \left(\sum_{j=1}^k ||D_j||^2 ||Q|| + ||M|| \right)\ dt \\ &= \frac{m^2}{-2\omega} \left(\sum_{j=1}^k ||D_j||^2 ||Q|| + ||M|| \right). \end{aligned}$$

By repeating this estimate n times, we obtain

$$||Q|| \leq \left(\frac{m^2 \sum_{j=1}^k ||D_j||^2}{-2\omega} \right)^n ||Q|| + \frac{m^2 ||M||}{-2\omega} \sum_{i=0}^{n-1} \left(\frac{m^2 \sum_{j=1}^k ||D_j||^2}{-2\omega} \right)^i.$$

Let $n \to \infty$ to obtain

$$||Q|| \leq \frac{m^2||M||}{-2\omega} \frac{1}{1 - \frac{m^2 \sum_{j=1}^k ||D_j||^2}{-2\omega}} = -\frac{m^2||M||}{m^2 \sum_{j=1}^k ||D_j||^2 + 2\omega}.$$

If $M = 0$ then $Q = 0$ by uniqueness of the solution.

Now suppose $M \leq 0$. Then we can check that the recursion for (Q_i) has the property that $Q_i \geq 0$ for all i. So $Q \geq 0$, and (6.29) shows that

$$Q \geq -\int_0^\infty T^*(t)MT(t)\ dt.$$

If $M < 0$, then there exists a unique $P \in L(H)$, $P > 0$ such that $M = L^*P + PL$. Then

$$Q \geq -\int_0^\infty T^*(t)MT(t)\ dt = P > 0.$$

□

So far we only know that if $Q \in L(H)$ a solution to (6.27) with $M < 0$, then $Q > 0$. But to obtain equivalence of norms we need Q to be coercive. In the finite-dimensional case coerciveness of Q is implied by $Q > 0$ but in infinite dimensions this is not the case. The next proposition shows that we can find a coercive solution in case L is dissipative, or equivalently if $(T(t))_{t\geq 0}$ is a contraction semigroup.

Proposition 6.21. *Suppose $L, (D_j)_{j=1,\dots,k}$ are as in Proposition 6.20 and that (6.26) holds. Then there exists a $Q \in L(H)$ such that (6.28) holds with $M = L + L^*$. Furthermore for this Q we have $Q \geq I$ and*

$$||Q|| \leq \frac{-2\omega}{-(2\omega + m^2 \sum_{j=1}^k ||D_j||^2)}.$$

Note that if L is dissipative, then $\langle Mx, x\rangle = 2\langle Lx, x\rangle \leq 0$ for all $x \in H$.

PROOF: By Proposition 6.20, there exists a unique solution $R \in L(H)$, $R \geq 0$ to (6.28) with $M = -\sum_{j=1}^k D_j^* D_j$. Furthermore $R \geq 0$ and, by (6.30)

$$||R|| \leq \frac{\sum_{j=1}^k ||D_j||^2 m^2}{-(2\omega + m^2 \sum_{j=1}^k ||D_j||^2)}.$$

Let $Q := I + R$. Then $Q \geq I$, the claimed estimate for Q holds and

$$L^*Q + QL + \sum_{j=1}^k D_j^* Q D_j = L^* + L + \sum_{j=1}^k D_j^* D_j - \sum_{j=1}^k D_j^* D_j = L^* + L,$$

interpreted in the weak sense of (6.28). □

6.6.2 Unbounded case

Recall that H be a real Hilbert space, and we consider the linear SDE in mild form

$$X(t) = S(t)x + \int_0^t S(t-s)BX(s)\; dW(s), \tag{6.31}$$

where W is a standard Brownian motion in $\mathbb{R}$, $(S(t))_{t\geq 0}$ is a strongly continuous semigroup on H with generator $A : \mathfrak{D}(A) \to H$, $B \in L(H)$ and $x \in H$.

Remark 6.22. From this point onward we assume for notational convenience $k = 1$, i.e. the stochastic process X is driven by only one standard Brownian motion. There is however no problem in proving all the results of this section for the case with multiple Brownian motions.

Lemma 6.23. *Suppose $T = (T(t))_{t\geq 0}$ is a strongly continuous semigroup with approximation $(T_n)_{n\in\mathbb{N}}$ such that, for some $\omega < 0$ and $m \geq 1$,*

$$||T(t)|| \vee \sup_{n\in\mathbb{N}} ||T_n(t)|| \leq m\mathrm{e}^{\omega t}.$$

For $n \in \mathbb{N}$ let $R_n \geq 0$ denote the unique positive semidefinite solution in $L(H)$ to the Lyapunov equation

$$A_n^* R_n + R_n A_n = M$$

for some fixed $M \in L(H)$, $M \leq 0$.

Then $\langle y, R_n x\rangle \to \langle y, Rx\rangle$ for any $x, y \in H$, where $R \geq 0$ is the unique positive semidefinite solution in $L(H)$ to

$$A^* R + RA = M.$$

Proof: We have, for any $t > 0$,

$$\begin{aligned}
&|\langle y, R_n x - Rx\rangle| \\
&= \left| \left\langle y, \int_0^\infty T_n^*(s) M T_n(s)x - T^*(s)MT(s)x\; ds \right\rangle \right| \\
&= \left| \int_0^\infty \langle T_n(s)y, M(T_n(s) - T(s))x\rangle + \langle (T_n(s) - T(s))y, MT(s)x\rangle\; ds \right| \\
&\leq \int_0^\infty ||T_n(s)||\; ||y||\; ||M||\; |(T_n(s) - T(s))x|\; ds \\
&\quad + \int_0^\infty |(T_n(s) - T(s))y|\; ||M||\; ||T(s)||\; ||x||\; ds \\
&\leq m||M|| \int_0^\infty \mathrm{e}^{\omega s} \left(||y||\; |(T_n(s) - T(s))x| + |(T_n(s) - T(s))y|\; ||x|| \right)\; ds.
\end{aligned}$$

For the first term we have

$$\int_0^\infty \mathrm{e}^{\omega s}\; |(T_n(s) - T(s))x| \leq \int_0^t \mathrm{e}^{\omega s} |(T_n(s) - T(s))x|\; ds + 2m|x| \int_t^\infty \mathrm{e}^{2\omega s}\; ds.$$

Now pick t large enough such that the second term is smaller than $\varepsilon/2$. Since (T_n) is an approximation of T, we have uniform convergence in $s \in [0,t]$ of $|T_n(s)x - T(s)x|$. So let N large enough such that $|T_n(s)x - T(s)x| \leq \delta$ for all $s \in [0,t]$ and for $\delta > 0$ such that

$$\int_0^t e^{\omega s} \delta \, ds < \varepsilon/2.$$

Repeating this argument for the second term leads to the stated result. □

Lemma 6.24. *Suppose $(T(t))_{t\geq 0}$ is a strongly continuous semigroup with approximation $(T_n(t))_{t\geq 0, n\in\mathbb{N}}$ and infinitesemal generators L and $(L_n)_{n\in\mathbb{N}}$, respectively. Suppose that for some $m \geq 1$ and $\omega < 0$ we have*

$$||T(t)|| \vee \sup_{n\in\mathbb{N}} ||T_n(t)|| \leq m e^{\omega t},$$

and suppose for this m, ω and some $D \in L(H)$ condition (6.26) holds.

Let $M \in L(H)$ be self-adjoint and negative semidefinite. Let Q and Q^n, $n \in \mathbb{N}$, denote the unique positive semidefinite solutions to

$$L^*Q + QL + D^*QD = M \quad \text{and} \quad L_n^*Q^n + Q^n L_n + D^*Q^n D = M.$$

Then for all $x, y \in H$ we have that $\langle x, Q^n y\rangle \to \langle x, Qy\rangle$ as $n \to \infty$.

Proof: For $n \in \mathbb{N}$ construct a recursion by

$$Q_0^n := 0 \quad \text{and} \quad Q_{j+1}^n := \int_0^\infty T_n^*(t)(D^*Q_j^n D - M)T_n(t) \, dt, \quad j \in \mathbb{N}.$$

Similarly let

$$Q_0 := 0 \quad \text{and} \quad Q_{j+1} := \int_0^\infty T^*(t)(D^*Q_j D - M)T(t) \, dt, \quad j \in \mathbb{N}.$$

First we will prove the following

Claim: For all $x, y \in H$ and $j \in \mathbb{N}$ we have $\langle x, (Q_j^n - Q_j)y\rangle \to 0$ as $n \to \infty$.

Proof of claim: For $j = 0$ the claim holds trivially. Suppose now that the claim holds for value $j = k-1$. Let $x, y \in H$. Then for $j = k$,

$$\begin{aligned}
|\langle x, (Q_j^n - Q_j)y\rangle| &= \left| \int_0^\infty \langle T_n(t)x, (D^* Q_{j-1}^n D - M)T_n(t)y\rangle \; dt \right. \\
&\qquad \left. - \int_0^\infty \langle T(t)x, (D^* Q_{j-1} D - M)T(t)y\rangle \; dt \right| \\
&\leq \int_0^\infty |\langle T_n(t)x, D^*(Q_{j-1}^n - Q_{j-1})DT_n(t)y\rangle| \; dt \\
&\quad + \left| \left\langle x, \int_0^\infty T_n^*(t)(D^* Q_{j-1} D - M)T_n(t)dt \; y \right\rangle \right. \\
&\qquad \left. - \left\langle x, \int_0^\infty T(t)^*(D^* Q_{j-1} D - M)T(t) \; dt \; y \right\rangle \right|
\end{aligned}$$

Since by the induction hypothesis $Q_{j-1}^n \to Q_{j-1}$ in weak sense, we have that

$$\langle DT_n(t)x, (Q_{j-1}^n - Q_{j-1})DT_n(t)y\rangle \to 0 \quad \text{as } n \to \infty$$

for all $t \in [0, \infty)$. By dominated convergence therefore the first term proceeds zero as $n \to \infty$. The convergence of the second term is an immediate consequence of Lemma 6.23.

So the claim is proven by induction. $\diamond$

Now by the proof of Theorem 6.20, $Q_j \to Q$ and $Q_j^n \to Q^n$ in the norm topology of $L(H)$, uniformly in n. Therefore using

$$|\langle x, (Q^n - Q)y\rangle| \leq |\langle x, (Q^n - Q_j^n)y\rangle| + |\langle x, (Q_j^n - Q_j)y\rangle| + |\langle x, (Q_j - Q)y\rangle|$$

we obtain $\langle x, (Q^n - Q)y\rangle \to 0$ as $n \to \infty$. $\square$

Lemma 6.25. *Let $B \in L(H)$ and let $k \in \mathbb{R}$ such that $B - kI$ is stable, i.e. we may estimate $||e^{-kt}e^{Bt}|| \leq me^{\lambda t}$, $t \geq 0$ for some $m \geq 1$ and $\lambda < 0$.*

Then for any $Q \in L(H)$, $Q \geq 0$, and $x \in H$ we have

$$\frac{\langle QBx, x\rangle}{\langle Qx, x\rangle} \leq k.$$

Proof: Let

$$N := (B - kI)^* Q + Q(B - kI).$$

Then by Lyapunov theory $N \leq 0$. Hence

$$2\langle Q(B - kI)x, x\rangle = \langle Nx, x\rangle \leq 0$$

or equivalently

$$\langle QBx, x\rangle \leq k\langle Qx, x\rangle.$$

$\square$

We are now ready to state and prove the main result of this section.

Theorem 6.26. *Let A be the infinitesemal generator of a strongly continuous semigroup in $L(H)$ and let $B \in L(H)$. Suppose there exists $b \in \mathbb{R}$ such that the semigroup $S(t)$ generated by $A + bB$ satisfies, for some $\mu \in \mathbb{R}$,*

$$||S(t)|| \leq \mathrm{e}^{\mu t}, \quad t \geq 0.$$

Furthermore suppose that for some $\lambda \in \mathbb{R}$,

$$2(\mu + \tfrac{1}{4}b^2 - \lambda) + ||B||^2 < 0.$$

Then

$$\limsup_{t \to \infty} \frac{1}{t} \log |X(t)| \leq \lambda,$$

where $(X(t))_{t \geq 0}$ is the mild solution of

$$dX(t) = AX(t)\ dt + BX(t)\ dW(t), \quad X(0) = x,$$

with W a one-dimensional standard Brownian motion.

Proof: Define $L := A + bB + (\frac{1}{4}b^2 - \lambda)I$, then L is the generator of a semigroup $T(t)$ and we have

$$||T(t)|| \leq \mathrm{e}^{(\mu + \frac{1}{4}b^2 - \lambda)t}.$$

In particular, T is a contraction semigroup and hence also the approximating semigroups $(T_n)_{n \in \mathbb{N}}$ (generated by the Yosida approximation $(L_n)_{n \in \mathbb{N}}$) are contraction semigroups. Furthermore, by (6.16), for any $\varepsilon > 0$ there exists an $N \in \mathbb{N}$ such that for all $n \geq N$ we have

$$||T_n(t)|| \leq \mathrm{e}^{(\mu + \frac{1}{4}b^2 - \lambda + \varepsilon)t}.$$

Let $\varepsilon > 0$, small enough such that $\mu + \frac{1}{4}b^2 - \lambda + \frac{1}{2}||B||^2 + \varepsilon < 0$ and let N as above. Let $\omega := \mu + \frac{1}{4}b^2 - \lambda + \varepsilon$. For $n \geq N$ let Q_n be the solution given by Proposition 6.21 to

$$L_n^* Q_n + Q_n L_n + B^* Q_n B = L_n^* + L_n.$$

Similarly let Q be the solution to

$$L^* Q + QL + B^* QB = L^* + L.$$

Recall that $Q_n \geq I$, $n \geq N$, $Q \geq I$ and

$$||Q|| \vee \sup_{n \geq N} ||Q_n|| \leq \frac{-2\omega}{-(2\omega + ||B||^2)}.$$

Let X_n denote the solution to

$$dX_n(t) = A_n X_n(t)\ dt + BX_n(t)\ dW(t), \quad X_n(0) = x,$$

where $A_n := L_n - bB - (\frac{1}{4}b^2 - \lambda)I$. Then (A_n) is an approximation for A and hence by Proposition 6.14, we have almost surely that, for a subsequence $(n_k) \subset \mathbb{N}$,

$$\sup_{t\in[0,T]} |X_{n_k}(t) - X(t)| \to 0 \quad \text{for all } T > 0.$$

Since L_n is bounded, we have that, since

$$(A_n + bB)^* Q_n + Q_n(A_n + bB) + B^* Q_n B + (\tfrac{1}{2}b^2 - 2\lambda)Q_n \leq 0, \quad n \in \mathbb{N},$$

and by the proofs of Lemma 6.18 and Proposition 6.19 that

$$\log\langle Q_n X_n(t), X_n(t)\rangle \leq \log\langle Q_n x, x\rangle + 2\lambda t + 2\int_0^t \frac{\langle Q_n B X_n(s), X_n(s)\rangle}{\langle Q_n X_n(s), X_n(s)\rangle}\, dW(s). \quad (6.32)$$

Here we used that $L_n + L_n^* \leq 0$ by dissipativeness of L_n.

In the stochastic integral we have by Lemma 6.24, the almost sure convergence of $X_n(s)$ on $[0,t]$ and the uniform boundedness of $||Q_n||$, $n \in \mathbb{N}$, that the integrand converges almost surely. By Lemma 6.25 we may apply dominated convergence to have convergence in $L^2(\Omega)$ of the stochastic integral

$$\int_0^t \frac{\langle Q_n B X_n(s), X_n(s)\rangle}{\langle Q_n X_n(s), X_n(s)\rangle}\, dW(s) \to \int_0^t \frac{\langle QBX(s), X(s)\rangle}{\langle QX(s), X(s)\rangle}\, dW(s).$$

We therefore also have this convergence with probability one for a further subsequence $(n_{k_l}) \subset \mathbb{N}$.

For this subsequence the lefthand side of (6.32) converges a.s. to $\log\langle QX(t), X(t)\rangle$, and by the uniform estimate on $||Q_n||$, the term $\log\langle Q_n x, x\rangle$ is bounded by a constant, say $M > 0$. Hence we have that, with probability one,

$$\log\langle QX(t), X(t)\rangle \leq M + 2\lambda t + 2\int_0^t \frac{\langle QBX(s), X(s)\rangle}{\langle QX(s), X(s)\rangle}\, dW(s), \quad \text{for all } t \geq 0.$$

Dividing by t and by letting $t \to \infty$, using the law of large numbers for martingales (Theorem 6.11) we obtain

$$\limsup_{t\to\infty} \frac{1}{t} \log\langle QX(t), X(t)\rangle \leq 2\lambda \quad \text{almost surely.}$$

Using the fact that $Q \geq I$, we now have

$$\limsup_{t\to\infty} \frac{1}{t} \log|X(t)| \leq \lambda \quad \text{almost surely.}$$

□

6.6.3 Example: stochastic delay differential equation

Consider the stochastic differential equation with delay

$$dY(t) = aY(t) + cY(t-1)\ dt + \sigma Y(t)\ dW(t), \quad t \geq 0, \quad Y(0) = y. \tag{6.33}$$

with $a, c \in \mathbb{R}$ and $\sigma > 0$.

Suppose first $c = 0$. Then the solution to the stochastic differential equation is given by

$$Y(t) = \exp((a - \tfrac{1}{2}\sigma^2)t + \sigma W(t))y, \quad t \geq 0,$$

and the solution is pathwise asymptotically stable if $a < \frac{1}{2}\sigma^2$.

We may now ask ourselves the question: for which $c \in \mathbb{R}$ do we still have stability?

Proposition 6.27. *Suppose $a < \frac{1}{2}\sigma^2$ and*

$$|c| < \mathrm{e}^{-3/2\sigma^2}(\tfrac{1}{2}\sigma^2 - a). \tag{6.34}$$

Then the solution to (6.33) *is pathwise exponentially stable.*

Proof: As in Section 3.1, let A be the generator of the delay semigroup in $\mathbb{R} \times L^2([-1,0])$, with

$$A := \begin{bmatrix} a & \Phi \\ 0 & \frac{d}{d\sigma} \end{bmatrix},$$

where $\Phi u := cu(-1)$ for $u \in W^{1,2}([-1,0])$. Furthermore let $B \in L(\mathbb{R}; \mathbb{R} \times L^2([-1,0]))$ be given by

$$B := \begin{bmatrix} \sigma & 0 \\ 0 & 0 \end{bmatrix}.$$

By Theorem 3.18, we have that if, for some $b \in \mathbb{R}$,

$$(a + b\sigma - \mu)^2 > c^2 \mathrm{e}^{-2\mu} \tag{6.35}$$

and

$$\mu > a + b\sigma, \tag{6.36}$$

then $A + bB - \mu I$ generates a dissipative semigroup on a renormed space $\mathbb{R} \times L^2([-1,0],\tau)$ (with $L^2([-1,0],\tau)$ the Hilbert space consisting of square integrable functions on $[0,1]$ with inner product

$$\langle f, g \rangle_\tau = \int_{-1}^{0} f(s) g(s) \tau(s)\ ds$$

for some suitable weight function $\tau \in L^\infty([-1,0])$.

By Theorem 6.26, we have that (6.33) has a pathwise exponentially stable solution if

$$\sigma^2 + \tfrac{1}{2}b^2 + 2\mu < 0. \tag{6.37}$$

We may reformulate (6.35) as

$$c^2 \leq (a + b\sigma - \mu)^2 e^{2\mu},$$

where b and μ should satisfy (6.36) and (6.37). It may be verified that $b = -2\sigma$ and $\mu = -3/2\sigma^2 - \varepsilon$, with $\varepsilon > 0$ sufficiently small, satisfy these conditions, using that $a < \frac{1}{2}\sigma^2$. By letting $\varepsilon \downarrow 0$ we obtain the estimate (6.34). □

Remark 6.28. From the proof of Proposition 6.27, we see that the best estimate for $|c|$ in order for the system to remain asymptotically stable is obtained by solving the nonlinear optimization problem

$$\begin{aligned} \max \quad & (a + b\sigma - \mu)^2 e^{2\mu} \\ \text{subject to} \quad & a + b\sigma - \mu < 0 \\ \text{and} \quad & \sigma^2 + \tfrac{1}{2}b^2 + 2\mu < 0. \end{aligned} \tag{6.38}$$

over b and μ. The condition $a < \frac{1}{2}\sigma^2$ is required for the set of feasible (b, σ) to be non-empty. The above problem may be solved by applying the Karush-Kuhn-Tucker conditions, obtaining $\mu = -\frac{1}{4}b^2 - \frac{1}{2}\sigma^2$ and for b the third-degree equation

$$\tfrac{1}{4}b^3 + b^2\sigma + b(a + \tfrac{1}{2}\sigma^2 - 1) - 2\sigma = 0.$$

By solving this equation we obtain an estimate which is sharper than (6.34), but less readable. ◇

Remark 6.29. Note that, for $\sigma = 0$, we obtain from (6.34) the condition $|c| < -a$, which is the same estimate as that in Corollary 3.19. Furthermore we have

$$\frac{d}{d(\sigma^2)} e^{-3/2\sigma^2}(\tfrac{1}{2}\sigma^2 - a)\Big|_{\sigma^2=0} = \tfrac{1}{2}(3a + 1),$$

from which we may conclude (see (6.34)) that adding noise has a stabilizing effect for $a > -\frac{1}{3}$. ◇

Example 6.30 (Population growth under random migration). Consider the example described in Chapter 1 and more explicitly in Example 3.3.1, with equation

$$dx(t) = [-\alpha x(t) + \beta x(t-1)]\, dt + \sigma x(t)\, dW(t), \quad t \geq 0,$$

so in (6.33) we have $a = -\alpha$ and $c = \beta$. Then Proposition 6.27 tells us that for $\beta < e^{-3/2\sigma^2}(\frac{1}{2}\sigma^2 + \alpha)$ the population will eventually be extinguished with probability one.

If, for example $\alpha = 0.1$, then the graph in Figure 6.1 shows upper bounds on values of β for which we know to have pathwise stability.

Numerical experiments suggest that for σ not too large the theoretical bounds are quite accurate (see Figure 1.1 in Chapter 1).

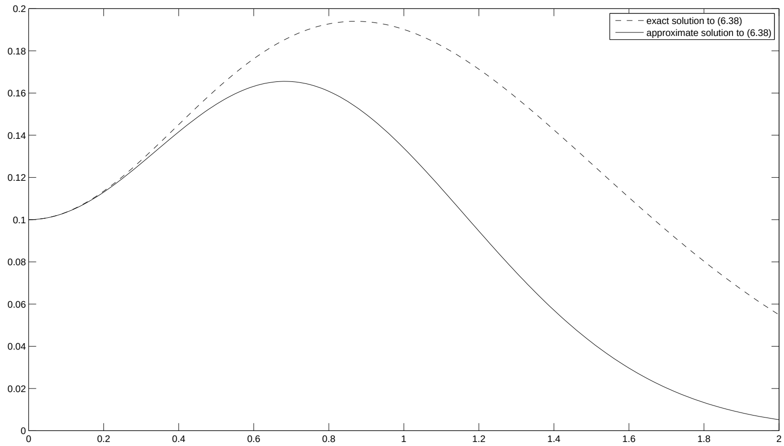


Figure 6.1: Bound on the birth rate β for which we have extinction of the population, as function of σ. Here $\alpha = 0.1$. The solid graph represents the approximation $\exp(-3/2\sigma^2)(\frac{1}{2}\sigma^2 + \alpha)$ provided by Proposition 6.27. The dashed graph represent the exact solution to the optimization problem (6.38), which gives a more relaxed requirement on β.

6.6.4 Relation to moment stability

So far we studied pathwise stability properties of stochastic differential equations. We will now relate these results to second moment stability for linear stochastic differential equations with multiplicative noise. Again let $(X(t;x))_{t\geq 0}$ be the solution to (6.23).

We recall the following result ([24], Theorem 11.14).

Theorem 6.31. *The following statements are equivalent.*

(i) There exist $M > 0$, $\omega > 0$ *such that*

$$\mathbb{E}|X(t;x)|^2 \leq M\mathrm{e}^{-\omega t}|x|^2, \quad t \geq 0, x \in H.$$

(ii) For any $x \in H$ *we have*

$$\mathbb{E}\int_0^\infty |X(t;x)|^2 \, dt < \infty.$$

(iii) There exists a positive semi-definite solution $R \in L(H)$ *to the equation*

$$A^*R + RA + \sum_{j=1}^{k} B_j^* R B_j + I = 0.$$

Combining the above theorem with Theorem 6.20, we obtain the following result:

Corollary 6.32. *Let* A *generate a strongly continuous semigroup* $(T(t))_{t\geq 0}$ *satisfying*

$$||T(t)|| \leq m\mathrm{e}^{\lambda t}, \quad t \geq 0$$

for some $m \geq 1$ *and* $\lambda < 0$. *Furthermore suppose that*

$$2\lambda + m^2 \sum_{j=1}^{k} ||B_j||^2 < 0.$$

Then there exist $M > 0$ *and* $\omega > 0$ *such that*

$$\mathbb{E}|X(t;x)|^2 \leq M\mathrm{e}^{-\omega t}|x|^2, \quad t \geq 0, x \in H.$$

This result can also be obtained directly from the estimate

$$\mathbb{E}|\mathrm{e}^{-\lambda t}X(t;x)|^2 \leq m^2|x|^2 + m^2 \int_0^t \sum_{j=1}^{k} ||B_j||^2 \mathbb{E}|\mathrm{e}^{-\lambda s}X(s;x)|^2 \, ds$$

and the Gronwall inequality.

In particular, it follows that under the conditions of Theorem 6.26, the process $Z(t;x)$ is stable in second moment, where $Z(t;x)$ is the solution of

$$Z(t;x) = S(t)x + \int_0^t S(t-s)BZ(s;x)\, dW(s),$$

with $(S(t))_{t\geq 0}$ the strongly continuous contraction semigroup generated by $A+bB+(\frac{1}{4}b^2-\lambda)I$.

6.7 Notes and remarks

Part of this chapter has been submitted for publication. A paper based on Section 6.6 will be submitted in the autumn of 2009.

For more results on finite dimensional stochastic Lyapunov exponents, see the overview paper [6], the book by Khasminskii [47], and e.g. [4], [5], [51], [69] and [77].

Results on pathwise stability with general decay rate (i.e. not necessarily exponential decay) of stochastic evolutions are given in [14]. In [15] results on stabilization by noise of some partial differential equations may be found. In [17] results on stabilization by noise for stochastic reaction-diffusion equations are used to establish existence and uniqueness of invariant measure.

In [52] a result on pathwise stability of finite dimensional stochastic differential equations with jumps is established, using the existence of an invariant measure for the projection of the solution on the unit sphere.

There exists a large amount of literature on stability of stochastic delay equations, of which we mention [1], [2], [3], [57], [64], [65].

Based on the contents of Section 6.6, a paper has been submitted for publication ([11], containing the bounded case), and another paper is being prepared for publication ([9], including the unbounded case).

APPENDIX

A

Nuclear and Hilbert-Schmidt operators

The following results are standard and can be found in e.g. [24], Appendix C. The only difference in our presentation is that we allow Hilbert-Schmidt operators to map into non-separable Hilbert spaces (see Proposition A.1).

See also [30], Section XI.6, for information on Hilbert-Schmidt operators. For an account of nuclear operators and Hilbert-Schmidt operators mapping one space into another, see [72].

Let X, Y be Banach spaces. An element $T \in L(X;Y)$ is said to be a *nuclear operator* if there exists two sequences $(y_i) \subset Y$, $(\varphi_i) \subset X^*$ such that

$$\sum_{i=1}^{\infty} ||y_i|| \, ||\varphi_i|| < \infty$$

and T has the representation

$$Tx = \sum_{i=1}^{\infty} y_i \varphi_i(x), \quad x \in X.$$

The space of all nuclear operators from X into Y, equipped with the norm

$$||T||_1 = \inf \left\{ \sum_{i=1}^{\infty} ||y_i|| \, ||\varphi_i|| : Tx = \sum_{i=1}^{\infty} y_i \varphi_i(x) \right\}$$

is a Banach space and will be denoted by $L_1(X;Y)$. As usual we let $L_1(X)$ denote $L_1(X;X)$.

Let E be another Banach space. If $T \in L_1(X;Y)$ and $S \in L(Y;E)$ then $TS \in L_1(X;E)$ and $||TS||_1 \leq ||T|| \, ||S||_1$.

Let $(H, \langle \cdot, \cdot \rangle, |\cdot|)$ be a separable Hilbert space and let (e_i) be a complete orthonormal system in H. If $T \in L_1(H)$ then we define the trace of T, independent of the orthonormal system (e_i) by

$$\text{tr } T = \sum_{i=1}^{\infty} \langle Te_i, e_i \rangle.$$

Furthermore, for $T \in L_1(H)$ and $S \in L(H)$ we have $TS, ST \in L_1(H)$ and

$$\text{tr } TS = \text{tr } ST \leq ||T||_1 ||S||.$$

A self-adjoint, nonnegative operator $T \in L(H)$ is nuclear if and only if for a complete orthonormal system (e_i) on H we have

$$\sum_{i=1}^{\infty} \langle Te_i, e_i \rangle < \infty.$$

In this case tr $T = ||T||_1$.

Let H, K be two Hilbert spaces, of which H is separable with complete orthonormal system (e_i). An operator $T \in L(H; K)$ is said to be *Hilbert-Schmidt* if

$$\sum_{i=1}^{\infty} |Te_i|^2 < \infty.$$

Proposition A.1. *The definition of Hilbert-Schmidt operator and of the number*

$$||T||_{L_{\text{HS}}} := \left(\sum_{i=1}^{\infty} |Te_i|^2 \right)^{1/2}$$

is independent of the choice of basis (e_i).

Proof: Since we can represent any $T \in L(H; K)$ as

$$Tx = \sum_{i=1}^{\infty} \langle e_i, x \rangle k_i, \quad x \in H$$

for some sequence $(k_i) \subset K$, we see that T is separably valued, i.e. there exists a closed linear subspace $M \subset K$ which is separable and such that $T \in L(H; M)$. Now proceed as in [24], Appendix C. □

Remark A.2. In [24], in the definition of Hilbert-Schmidt operator K is also assumed to be separable. Our line of reasoning in the proof of the above proposition shows that this assumption on K is not necessary.

Proposition A.3. *Suppose* $T \in L_{\mathrm{HS}}(H;K)$, $S \in L(K;M)$ *and* $U \in L(F;H)$ *where* F, H, K *and* M *are Hilbert spaces, of which* H *and* F *are separable. Then* $ST \in L_{\mathrm{HS}}(H;M)$ *and* $TU \in L_{\mathrm{HS}}(F;K)$, *with*

$$||ST||_{L_{\mathrm{HS}}(H;M)} \leq ||S||_{L(K;M)}||T||_{L_{\mathrm{HS}}(H;K)} \quad \textit{and}$$
$$||TU||_{L_{\mathrm{HS}}(F;K)} \leq ||T||_{L_{\mathrm{HS}}(H;K)}||U||_{L(F;H)}.$$

PROOF: Let (e_i) and (f_j) be complete orthonormal systems in H and F, respectively. The first statement is trivial:

$$\sum_{i=1}^{\infty} |STe_i|^2 \leq ||S||^2 \sum_{i=1}^{\infty} |Te_i|^2.$$

To prove the second statement, we calculate

$$\begin{aligned} ||TU||_{L_{\mathrm{HS}}(F;K)} &= \sum_{j=1}^{\infty} |TUf_j|^2 = \sum_{j=1}^{\infty}\sum_{i=1}^{\infty} |Te_i\langle Uf_j, e_i\rangle_H|^2 \\ &= \sum_{j=1}^{\infty}\sum_{i=1}^{\infty} |Te_i|^2\langle Uf_j, e_i\rangle_H^2 = \sum_{i=1}^{\infty} |Te_i|^2 \sum_{j=1}^{\infty}\langle f_j, U^*e_i\rangle_H^2 \\ &= \sum_{i=1}^{\infty} |Te_i|^2|U^*e_i|^2 \leq ||U^*||^2 \sum_{i=1}^{\infty} |Te_i|^2 = ||U||^2||T||_{L_{\mathrm{HS}}(H;K)}. \end{aligned}$$

where the manipulations are allowed by absolute convergence. □

Proposition A.4. *Suppose* $S \in L_{\mathrm{HS}(K;H)}$ *and* $T \in L_{\mathrm{HS}(H;K)}$, *with* H *and* K *separable Hilbert spaces. Then* $ST \in L_1(H)$, $TS \in L_1(K)$ *and* $\mathrm{tr}\ ST = \mathrm{tr}\ TS$.

PROOF: Let (e_i) be a complete orthonormal system in H and (f_j) in K.

$$\begin{aligned} \mathrm{tr}\ ST &= \sum_{i=1}^{\infty}\langle STe_i, e_i\rangle = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\langle Te_i, f_j\rangle\langle Sf_j, e_i\rangle \\ &= \sum_{j=1}^{\infty}\sum_{i=1}^{\infty}\langle Te_i, f_j\rangle\langle Sf_j, e_i\rangle = \mathrm{tr}\ TS, \end{aligned}$$

where the change of order in limit evaluation is allowed since

$$\begin{aligned} \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} |\langle Te_i, f_j\rangle\langle Sf_j, e_i\rangle| &\leq \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} \tfrac{1}{2}\left(\langle Te_i, f_j\rangle^2 + \langle Sf_j, e_i\rangle^2\right) \\ &= \tfrac{1}{2}\left(||T||^2_{L_{\mathrm{HS}(H;K)}} + ||S||^2_{L_{\mathrm{HS}(K;H)}}\right). \end{aligned}$$

□

Let $L_{\mathrm{HS}}(H;K)$ denote the space of Hilbert-Schmidt operators from H into K. We can equip $L_{\mathrm{HS}}(H;K)$ with an inner product defined by

$$\langle S,T\rangle_{L_{\mathrm{HS}}(H;K)} = \sum_{i=1}^{\infty} \langle Se_i, Te_i\rangle.$$

Then $L_{\mathrm{HS}}(H;K)$ is a Hilbert space. In case both H and K are separable, $L_{\mathrm{HS}}(H;K)$ is separable with complete orthonormal system $(f_j \otimes e_i)$, where (f_j) is a complete orthonormal system in K, and where $x \otimes y$ denotes the linear operator $z \mapsto x\langle y,z\rangle \in L_{\mathrm{HS}}(H;K)$ for $x \in K$ and $y \in H$.

Remark A.5. Using similar arguments as used above, we can show the following:

(i) If H,K,M are Hilbert spaces of which H and K are separable, and if $S \in L_{\mathrm{HS}}(K;M)$ and $T \in L_{\mathrm{HS}}(H;K)$, then $ST \in L_1(H;M)$ and

$$||ST||_1 \leq ||S||_{L_{\mathrm{HS}}} ||T||_{L_{\mathrm{HS}}}.$$

(ii) If H,K are separable Hilbert spaces and $T \in L_{\mathrm{HS}}(H;K)$ then $T^* \in L_{\mathrm{HS}}(K;H)$.

Notation

Symbol	**Explanation**	**Refer to**
δ	Skorohod integral	Section 5.2.2
$\rho(A)$	resolvent set of A	Section 6.1.1
$\sigma(A)$	spectrum of A	Section 6.1.1
$\omega_0(A)$	growth bound of A	Section 6.1.1
Ω	probability space	[95]
A_n	Yosida approximation of A	Section 6.4.2
$\mathcal{B}(V)$	Borel σ-algebra of a topological space V	[95]
$C(U;V)$	continuous functions mapping U into V	[66]
$\mathfrak{D}(A)$	domain of the linear operator A	[76]
dF	Fréchet derivative of a mapping F	[40]
DX	Malliavin derivative of random variable X	Section 5.2.2
$\mathbb{E}$	Expectation operator corresponding to $\mathbb{P}$	[95]
$\mathcal{E}^p$	canonical state space for delay equations	Section 3.1
$[F]_{\mathrm{Lip}}$	Lipschitz constant of F	[96]
$\mathcal{F}$	σ-algebra on Ω	[95]
$(\mathcal{F}_t)_{t\geq 0}$	filtration on Ω	[73]
$\mathbb{H}(K)$	domain of Malliavin derivative operator of K-valued random variables	Section 5.2.2
$\mathcal{L}(X)$	probability law of a random variable X	[95]
$L(U)$	bounded linear operators mapping U into U	[76]
$L(U;V)$	bounded linear operators mapping U into V	[76]
$L_1(U)$	nuclear operators mapping U into U	Appendix A
$L_1^+(U)$	self-adjoint non-negative nuclear operators	Appendix A
$L_{\mathrm{HS}}(U;H)$	Hilbert-Schmidt operators mapping U into H	Appendix A
$L^2_{M,T}(H)$	space of H-valued processes which are integrable with respect to $M \in \mathcal{M}^2(U)$ on $[0,T]$	Section 2.2
$\mathcal{M}^2(U)$	square integrable U-valued martingales	Section 2.1
$N_{m,Q}$	normal distribution with mean m and covariance operator Q	[13], [24], [26]

Symbol	Explanation	Refer to
$\mathbb{P}$	probability measure on Ω	[95]
$\mathcal{P}_T$	predictable σ-algebra on $[0,T]\times\Omega$	[41]
$\mathfrak{r}(A)$	spectral radius of A	Section 6.1.1
RKHS	reproducing kernel Hilbert space	Section 2.3.3
$\mathfrak{s}(A)$	spectral bound of A	Section 6.1.1
tr T	trace of T	Appendix A
$W^{1,p}(U;V)$	Sobolev space	Section 3.1
$\langle\langle M\rangle\rangle$	operator angle bracket	Section 2.1
$\widehat{L}$	compensated version of a Lévy process L	Section 2.1.2
$U\hookrightarrow V$	the continuous injection of U into V	
$\frac{\partial u}{\partial \nu}$	derivative in the direction of the outward normal	[81]
T^*	adjoint operator of T	[76]

BIBLIOGRAPHY

[1] J. A. D. Appleby. Stabilisation of functional differential equations by noise. *Systems Control Lett*, 2003.

[2] J. A. D. Appleby and X. Mao. Stochastic stabilisation of functional differential equations. *Systems Control Lett.*, 54(11):1069–1081, 2005.

[3] J. A. D. Appleby and M. Riedle. Almost sure asymptotic stability of stochastic Volterra integro-differential equations with fading perturbations. *Stoch. Anal. Appl.*, 24(4):813–826, 2006.

[4] L. Arnold. A formula connecting sample and moment stability of linear stochastic systems. *SIAM J. Appl. Math.*, 44(4):793–802, 1984.

[5] L. Arnold, E. Oeljeklaus, and E. Pardoux. Almost sure and moment stability for linear Itô equations. *Lyapunov exponents, Lecture Notes in Mathematics*, 1186:129–159, 1985.

[6] L. Arnold and V. Wihstutz. Lyapunov exponents: a survey. In *Lyapunov exponents (Bremen, 1984)*, volume 1186 of *Lecture Notes in Math.*, pages 1–26. Springer, Berlin, 1986.

[7] A. Bátkai and S. Piazzera. *Semigroups for Delay Equations.* AK Peters, Ltd., 2005.

[8] J. Bierkens. Dissipativeness of the Delay Semigroup. *In preparation*, 2009.

[9] J. Bierkens. Estimates on the pathwise Lyapunov exponent for linear stochastic evolutions with multiplicative noise. *In preparation*, 2009.

[10] J. Bierkens. Sufficient conditions for the eventual strong Feller property for degenerate evolutions. *In preparation*, 2009.

[11] J. Bierkens, O. van Gaans, and S.M. Verduyn Lunel. Estimates on the pathwise Lyapunov exponent for linear stochastic differential equations with multiplicative noise. *Submitted*, 2008.

[12] J. Bierkens, O. van Gaans, and S.M. Verduyn Lunel. Existence of an invariant measure for stochastic evolutions driven by an eventually compact semigroup. *Journal of Evolution Equations*, 9(4):771–786, 2009.

[13] V.I. Bogachev. *Gaussian Measures.* American Mathematical Society, 1998.

[14] T. Caraballo, M. J. Garrido-Atienza, and J. Real. Asymptotic stability of nonlinear stochastic evolution equations. *Stochastic Anal. Appl.*, 21(2):301–327, 2003.

[15] T. Caraballo, K. Liu, and X. Mao. On stabilization of partial differential equations by noise. *Nagoya Math. J.*, 161:155–170, 2001.

[16] R.A. Carmona and M.R. Tehranchi. *Interest rate models: an infinite dimensional stochastic analysis perspective.* Springer Finance. Springer, Berlin, 2006.

[17] S. Cerrai. Stabilization by noise for a class of stochastic reaction-diffusion equations. *Probab. Theory Related Fields*, 133(2):190–214, 2005.

[18] A. Chojnowska-Michalik. Invariant densities for stochastic semilinear evolution equations and related properties of transition semigroups. In *Stochastic partial differential equations and applications (Trento, 2002)*, volume 227 of *Lecture Notes in Pure and Appl. Math.*, pages 105–120. Dekker, New York, 2002.

[19] A. Chojnowska-Michalik and B. Goldys. Existence, uniqueness and invariant measures for stochastic semilinear equations on Hilbert spaces. *Probability Theory and Related Fields*, 102(3):331–356, 1995.

[20] R.F. Curtain and H. Zwart. *An Introduction to Infinite-Dimensional Linear Systems Theory.* Springer, 1995.

[21] G. Da Prato. *An introduction to infinite-dimensional analysis. Universitext.* Springer, 2006.

[22] G. Da Prato, D. Gatarek, and J. Zabczyk. Invariant measures for semilinear stochastic equations. *Stochastic Analysis and Applications*, 10(4):387–408, 1992.

[23] G. Da Prato, S. Kwapień, and J. Zabczyk. Regularity of solutions of linear stochastic equations in Hilbert spaces. *Stochastics*, 23(1):1–23, 1987.

[24] G. Da Prato and J. Zabczyk. *Stochastic Equations in Infinite Dimensions.* Cambridge University Press, 1992.

[25] G. Da Prato and J. Zabczyk. *Ergodicity for Infinite Dimensional Systems.* Cambridge University Press, 1996.

[26] G. Da Prato and J. Zabczyk. *Second order partial differential equations in Hilbert spaces*, volume 293 of *London Mathematical Society Lecture Note Series.* Cambridge University Press, Cambridge, 2002.

[27] E. B. Davies. *Quantum theory of open systems.* Academic Press, London, 1976.

[28] M. C. Delfour. The largest class of hereditary systems defining a C_0 semigroup on the product space. *Canad. J. Math.*, 32(4):969–978, 1980.

[29] O. Diekmann, S.A. van Gils, S.M. Verduyn Lunel, and H.O. Walther. *Delay Equations: Functional-, Complex-, and Nonlinear Analysis.* Springer, 1995.

[30] N. Dunford and J. T. Schwartz. *Linear operators. Part II.* Wiley Classics Library. John Wiley & Sons Inc., New York, 1988.

[31] J.A. Dyson, R. Villella-Bressan, and G.F. Webb. Asynchronous exponential growth in an age structured population of proliferating and quiescent cells. *Mathematical Biosciences*, 177-178:73 – 83, 2002.

[32] K.J. Engel and R. Nagel. *One-Parameter Semigroups for Linear Evolution Equations.* Springer, 2000.

[33] M. Fuhrman. Smoothing properties of nonlinear stochastic equations in Hilbert spaces. *NoDEA Nonlinear Differential Equations Appl.*, 3(4):445–464, 1996.

[34] M. Fuhrman. On a class of stochastic equations in Hilbert spaces: solvability and smoothing properties. *Stochastic Anal. Appl.*, 17(1):43–69, 1999.

[35] A. A. Gushchin and U. Küchler. On stationary solutions of delay differential equations driven by a Lévy process. *Stochastic Process. Appl.*, 88(2):195–211, 2000.

[36] M. Hairer, J. C. Mattingly, and M. Scheutzow. Asymptotic coupling and a weak form of Harris' theorem with applications to stochastic delay equations, 2009. *Prob. Theo. Rel. Fields, to appear.*

[37] M. Hairer and J.C. Mattingly. Ergodicity of the 2D Navier-Stokes equations with degenerate stochastic forcing. *Annals of Mathematics*, 164:993–1032, 2006.

[38] J. Hale. *Theory of functional differential equations.* Springer, New York, second edition, 1977. Applied Mathematical Sciences, Vol. 3.

[39] E. Hausenblas and J. Seidler. Stochastic convolutions driven by martingales: maximal inequalities and exponential integrability. *Stoch. Anal. Appl.*, 26(1):98–119, 2008.

[40] E. Hille and R. S. Phillips. *Functional analysis and semi-groups.* American Mathematical Society, Providence, R. I., 1974.

[41] Jean Jacod and Albert N. Shiryaev. *Limit theorems for stochastic processes*, volume 288 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences].* Springer-Verlag, Berlin, second edition, 2003.

[42] O. Kallenberg. *Foundations of Modern Probability.* Springer, second edition, 2002.

[43] G. Kallianpur and J. Xiong. *Stochastic differential equations in infinite dimensional spaces.* Lecture Notes – Monograph Series. Institute of Mathematical Statistics, 1995.

[44] I. Karatzas and S.E. Shreve. *Brownian Motion and Stochastic Calculus.* Springer, 1991.

[45] R.Z. Khasminskii. Ergodic properties of recurrent diffusion processes and stabilization of the solution of the Cauchy problem for parabolic equations. *Theory of Probability and its Applications*, 5:196–214, 1960.

[46] R.Z. Khasminskii. Necessary and sufficient condition for the asymptotic stability of linear stochastic system. *Theory of Probability and its Applications*, 12:144–147, 1967.

[47] R.Z. Khasminskii. *Stochastic stability of differential equations*, volume 7 of *Monographs and Textbooks on Mechanics of Solids and Fluids: Mechanics and Analysis.* Sijthoff & Noordhoff, Alphen aan den Rijn, 1980. Translated from the Russian by D. Louvish.

[48] W. Kliemann. Recurrence and invariant measures for degenerate diffusions. *Ann. Probab.*, 15(2):690–707, 1987.

[49] A.A. Kwiecińska. Stabilization of evolution equations by noise. *Proc. Amer. Math. Soc.*, 130(10):3067–3074 (electronic), 2002.

[50] M.A. Leibowitz. Statistical Behavior of Linear Systems with Randomly Varying Parameters. *Journal of Mathematical Physics*, 4:852–858, 1963.

[51] A. Leizarowitz. Exact results for the Lyapunov exponents of certain linear Ito systems. *SIAM J. Appl. Math.*, 50(4):1156–1165, 1990.

[52] C. W. Li, Z. Dong, and R. Situ. Almost sure stability of linear stochastic differential equations with jumps. *Probab. Theory Related Fields*, 123(1):121–155, 2002.

[53] W.-T. Li, Z.-C. Wang, and J. Wu. Entire solutions in monostable reaction-diffusion equations with delayed nonlinearity. *J. Differential Equations*, 245(1):102–129, 2008.

[54] Y. Lv, W. Lv, and J. Sun. Convergence dynamics of stochastic reaction-diffusion recurrent neural networks in continuously distributed delays. *Nonlinear Anal. Real World Appl.*, 9(4):1590–1606, 2008.

[55] X. Mao. Stochastic stabilization and destabilization. *Systems & Control Letters*, 23(4):279–290, 1994.

[56] X. Mao. Stochastic Differential Equations and Applications. *Horwood, Chichester*, 1997.

[57] X. Mao and A. Shah. Exponential stability of stochastic differential delay equations. *Stochastics Stochastics Rep.*, 60(1-2):135–153, 1997.

[58] B. Maslowski and J. Seidler. Invariant measures for nonlinear SPDE's: uniqueness and stability. *Arch. Math. (Brno)*, 34(1):153–172, 1998. Equadiff 9 (Brno, 1997).

[59] B. Maslowski and J. Seidler. Probabilistic approach to the strong Feller property. *Probab. Theory Related Fields*, 118(2):187–210, 2000.

[60] B. Maslowski and J. Seidler. Strong Feller infinite-dimensional diffusions. In *Stochastic partial differential equations and applications (Trento, 2002)*, volume 227 of *Lecture Notes in Pure and Appl. Math.*, pages 373–387. Dekker, New York, 2002.

[61] M. Métivier and J. Pellaumail. *Stochastic integration.* Academic Press, New York, 1980.

[62] M. Métivier and G. Pistone. Une formule d'isométrie pour l'intégrale stochastique hilbertienne et équations d'évolution linéaires stochastiques. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 33(1):1–18, 1975.

[63] S.-E. A. Mohammed. Stochastic differential systems with memory: theory, examples and applications. In *Stochastic analysis and related topics, VI (Geilo, 1996)*, volume 42 of *Progr. Probab.*, pages 1–77. Birkhäuser Boston, Boston, MA, 1998.

[64] S.-E. A. Mohammed and M. K. R. Scheutzow. Lyapunov exponents of linear stochastic functional differential equations driven by semimartingales. I. The multiplicative ergodic theory. *Ann. Inst. H. Poincaré Probab. Statist.*, 32(1):69–105, 1996.

[65] S.-E. A. Mohammed and M. K. R. Scheutzow. Lyapunov exponents of linear stochastic functional-differential equations. II. Examples and case studies. *Ann. Probab.*, 25(3):1210–1240, 1997.

[66] J.R. Munkres. *Topology, Second Edition.* Prentice-Hall, 2000.

[67] D. Nualart. *The Malliavin Calculus And Related Topics.* Springer, second edition, 2006.

[68] A. W. Olbrot and L. Pandolfi. Null controllability of a class of functional-differential systems. *Internat. J. Control*, 47(1):193–208, 1988.

[69] É. Pardoux and V. Wihstutz. Lyapunov exponent and rotation number of two-dimensional linear stochastic systems with small diffusion. *SIAM J. Appl. Math.*, 48(2):442–457, 1988.

[70] S. Peszat and J. Zabczyk. Strong Feller property and irreducibility for diffusions on Hilbert spaces. *Ann. Probab.*, 23(1):157–172, 1995.

[71] S. Peszat and J. Zabczyk. *Stochastic partial differential equations with Lévy noise*, volume 113 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2007.

[72] A. Pietsch. *Operator ideals*, volume 20 of *North-Holland Mathematical Library*. North-Holland Publishing Co., Amsterdam, 1980.

[73] P.E. Protter. *Stochastic Integration and Differential Equations*. Springer, 2004.

[74] M. Reiß, M. Riedle, and O. van Gaans. Delay differential equations driven by Lévy processes: stationarity and Feller properties. *Stochastic Process. Appl.*, 116(10):1409–1432, 2006.

[75] L.C.G. Rogers and D. Williams. *Diffusions, Markov Processes and Martingales*, volume 2. J. Wiley and Sons, 1987.

[76] W. Rudin. *Functional analysis*. International Series in Pure and Applied Mathematics. McGraw-Hill Inc., New York, second edition, 1991.

[77] M. Scheutzow. Stabilization and destabilization by noise in the plane. *Stochastic Anal. Appl.*, 11(1):97–113, 1993.

[78] E. Schmidt. Über eine Klasse linearer funktionaler Differentialgleichungen. *Math. Ann.*, 70(4):499–524, 1911.

[79] A. N. Shiryaev. *Probability*, volume 95 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1996. Translated from the first (1980) Russian edition by R. P. Boas.

[80] E. D. Sontag. *Mathematical control theory*, volume 6 of *Texts in Applied Mathematics*. Springer, New York, second edition, 1998.

[81] M. E. Taylor. *Partial Differential Equations: Basic Theory*. Springer, 1996.

[82] C. C. Travis and G. F. Webb. Existence and stability for partial functional differential equations. In *Dynamical systems (Proc. Internat. Sympos., Brown Univ., Providence, R.I., 1974), Vol. II*, pages 147–151. Academic Press, New York, 1976.

[83] C. C. Travis and G. F. Webb. Existence, stability, and compactness in the α-norm for partial functional differential equations. *Trans. Amer. Math. Soc.*, 240:129–143, 1978.

[84] O. van Gaans and J. van Neerven. Invariant measures for stochastic Cauchy problems with asymptotically unstable drift semigroup. *Electron. Comm. Probab.*, 11:24–34 (electronic), 2006.

[85] O. van Gaans and S.M. Verduyn Lunel. Long Term Behavior Of Dichotomous Stochastic Differential Equations In Hilbert Spaces. *Communications in Contemporary Mathematics*, 6(3):349–376, 2004.

[86] O. A. van Herwaarden. Spread of pollution by dispersive groundwater flow. *SIAM J. Appl. Math.*, 54(1):26–41, 1994.

[87] J. van Neerven. Uniqueness of invariant measures for the stochastic Cauchy problem in Banach spaces. 127:491–517, 2001.

[88] J. van Neerven. *Stochastic Evolution Equations.* Internet Seminar 2008 Lecture Notes, 2008.

[89] J. van Neerven and M. Riedle. Solutions of stochastic delay equations in spaces of continuous functions.

[90] J. van Neerven and M. Riedle. A semigroup approach to stochastic delay equations in spaces of continuous functions. *Semigroup Forum*, 74(2):227–239, 2007.

[91] J. van Neerven, M. C. Veraar, and L. Weis. Stochastic evolution equations in UMD Banach spaces. *J. Funct. Anal.*, 255(4):940–993, 2008.

[92] J. van Neerven and L. Weis. Invariant measures for the linear stochastic Cauchy problem and R-boundedness of the resolvent. *J. Evol. Equ.*, 6(2):205–228, 2006.

[93] M. Veraar. *Stochastic Integration in Banach Spaces and Applications to Parabolic Evolution Equations.* PhD thesis, TU Delft, 2006.

[94] P. Weng and J. Wu. Wavefronts for a non-local reaction-diffusion population model with general distributive maturity. *IMA J. Appl. Math.*, 73(3):477–495, 2008.

[95] D. Williams. *Probability with martingales.* Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991.

[96] K. Yosida. *Functional analysis.* Classics in Mathematics. Springer-Verlag, Berlin, 1995. Reprint of the sixth (1980) edition.

INDEX

Samenvatting

Stochastische differentiaalvergelijkingen met tijdsvertraging vormen de inspiratie voor dit proefschrift. Voorbeelden van zulke vergelijkingen treden op bij populatie-modellen, regelsystemen met tijdsvertraging en ruis, lasers, economische modellen, neurale netwerken, milieuvervuiling en in vele andere modellen. In zulke modellen zijn we vaak geinteresseerd in de *evolutie* van een bepaalde grootheid, bijvoorbeeld de grootte van een populatie, of de hoeveelheid vervuiling op een bepaalde plek, veranderend in de tijd.

Een differentiaalvergelijking met tijdsvertraging, ofwel *delay vergelijking*, is een differentiaalvergelijking waarbij de verandering in de tijd van zo'n grootheid wordt uitgedrukt als functie van die grootheid op bepaalde tijdstippen in zowel heden als verleden. Dit in contrast met een gewone differentiaalvergelijking, waarbij de verandering in de tijd van een grootheid op een bepaald tijdstip slechts wordt uitgedrukt als functie van die grootheid op datzelfde tijdstip.

We kunnen voorts ook invloeden van onzekerheid of ruis toevoegen aan zo'n delay vergelijking. Vaak gebeurt dit door de verandering in de tijd van de grootheid ook te laten afhangen van een ruisproces zoals bijvoorbeeld een Brownse beweging of een Poisson proces. De zo verkregen differentiaalvergelijking met tijdsvertraging en onzekerheid wordt ook wel een stochastische differentiaalvergelijking met tijdsvertraging of *stochastische delay vergelijking* genoemd. Het veranderen in de tijd onder onzekerheid wordt *stochastische evolutie* genoemd. Waar we hieronder over differentiaalvergelijkingen spreken dient de lezer vooral aan dit soort vergelijkingen te denken.

Alleen het beschrijven van een model, ofwel, in wiskundige termen, het opstellen van een stelsel differentiaalvergelijkingen, geeft weinig voldoening. We willen uitspraken kunnen doen over eigenschappen van het model, ofwel het kwalitatieve gedrag van de oplossingen van het stelsel differentiaalvergelijkingen.

Het mooiste zou zijn als we voor een gegeven differentiaalvergelijking de expliciete oplossing kunnen bepalen. Dat is het volledige toekomstige gedrag van de door de vergelijking beschreven grootheid, als functie van de tijd en de bronnen van onzekerheid. Nu is dit bij de differentiaalvergelijkingen waarover dit werk gaat bijna altijd onmogelijk.

In plaats daarvan kan, voor een beter kwalitatief begrip van oplossingen, het probleem van het bestaan van *stationaire toestanden* bestudeerd worden. Een stationaire toestand is een toestand die, zodra deze zich voordoet, zich zal blijven herhalen. Denk bijvoorbeeld aan een constante waarde of een periodieke oplossing.

Dit brengt ons op het begrip *toestand*. Wat voor toestanden kan de oplossing van een stochastische delay differentiaalvergelijking aannemen, ofwel, wat is de *toestandsruimte* van zo'n vergelijking? Omdat bij een delay vergelijking de verandering in de tijd van de grootheid afhangt van het verleden, moeten we om de toekomst van de grootheid te kennen ook het verleden kennen. Daarom bestaat de *toestand* van de grootheid niet alleen uit zijn huidige waarde, maar ook van die waarden in het verleden die we nodig hebben voor het beschrijven van de toekomst de grootheid. Het gevolg hiervan is dat de toestandsruimte een functieruimte is, ofwel een oneindig dimensionale ruimte. Vanwege dit oneindig dimensionale karakter van delay vergelijkingen, hebben ook stochastische delay vergelijkingen een oneindig dimensionale toestandsruimte. Het wiskundige instrument dat de huidige toestand afbeeldt op toekomstige toestanden heet een *halfgroep*.

Er is al een ruime hoeveelheid theorie voorhanden op het gebied van oneindig dimensionale stochastische differentiaalvergelijkingen. Echter, stochastische delay vergelijkingen vallen door hun specifieke karakter vaak net buiten de bestaande theorie. Zo is typisch aan delay vergelijkingen dat oplossingen pas na verloop van tijd glad zijn (en niet meteen, zoals bij de warmtevergelijking uit de natuurkunde), en dat de ruis gedegenereerd is: de ruis beinvloedt de stochastische evolutie slechts in bepaalde richtingen en niet in alle richtingen zoals vaak voorkomt bij stochastische partiële differentiaalvergelijkingen. Dit heeft er mee te maken dat de ruis niet het verleden van de stochastische evolutie kan beinvloeden.

Dit proefschrift

Na de inleiding (Hoofdstuk 1) beschrijven we eerst wat oneindig dimensionale stochastische differentiaalvergelijkingen zijn (Hoofdstuk 2). In het daaropvolgende hoofdstuk (Hoofdstuk 3) gebruiken we deze theorie voor het precies omschrijven van stochastische delay vergelijkingen. Als ingrediënt hiervoor wordt de zogenaamde delay halfgroep gedefinieerd, die oplossingen beschrijft van lineaire delay vergelijkingen. In dat hoofdstuk worden ook enkele belangrijke eigenschappen van delay vergelijkingen omschreven: het op den duur compact zijn van de delay halfgroep, en het feit dat het inproduct op de toestandsruimte zo gekozen kan worden dat de delay halfgroep een gegeneraliseerde contractie is.

Dan kan het onderzoek naar het lange termijn gedrag van stochastische delay vergelijkingen beginnen. Hiervoor keren we terug bij het begrip stationaire toestand.

Voor stochastische differentiaalvergelijkingen is een stationaire toestand een kansverdeling op de ruimte van toestanden die invariant is onder de differentiaalvergelijking: als eenmaal deze kansverdeling is aangenomen hebben alle toekomstige waardes van de oplossing deze zelfde kansverdeling. Zo'n kansverdeling heet *invariante kansverdeling*.

In Hoofdstuk 4 wordt aangetoond dat, met behulp van het op den duur compact zijn van de delay halfgroep, het bestaan van een invariante kansverdeling kan worden vastgesteld onder redelijke voorwaarden.

In Hoofdstuk 5 worden met technieken uit de Malliavin calculus voorwaarden gegeven waarbij zo'n invariante kansmaat uniek is. Als dit het geval is zal, wegens het zogenaamde ergodisch principe, het gemiddelde lange termijn gedrag van oplossingen deze invariante kansverdeling aannemen.

Tenslotte wordt in Hoofdstuk 6 het lange termijn gedrag van lineaire stochastische differentiaalvergelijkingen met multiplicatieve ruis onderzocht. In het bijzonder worden voorwaarden gegeven voor het (padsgewijs) stabiel zijn van oplossingen van deze vergelijkingen. Eerst gebeurt dit voor het geval waarbij de ruis niet gedegenereerd is. Daarna bestuderen we het geval toegespitst op stochastische delay vergelijkingen, waarbij de ruis wel gedegenereerd is.

Curriculum Vitae

Joris Bierkens was born on July 6th, 1979 in Tilburg, The Netherlands. In 1984 Joris informed Sinterklaas that he would like to pursue a scientific career. In 1997 he completed his high school education at Mill-Hill College in Goirle. In that same year he started his combined studies in Applied Computer Science and Mathematics at Eindhoven University of Technology. As part of his M.Sc. program he performed research on acoustics at the Marcus Wallenberg Laboratory of the KTH in Stockholm in 2002. During his studies in Eindhoven he was active in organizational roles for the cultural student community and in the university council. Under the supervision of prof. dr. ir. J. de Graaf he wrote his Master's thesis entitled *Geometrical Methods in Diffusion Tensor Regularization*. On May 15th, 2004, Joris obtained his M.Sc. degree in Applied Mathematics cum laude. This enabled him to participate in the Part III Maths program at the University of Cambridge, UK, where he obtained the Certificate of Advanced Study in Mathematics. In december 2005 he started his Ph.D. research under the supervision of prof. dr. S.M. Verduyn Lunel and dr. ir. O. van Gaans, resulting in this thesis. Currently he is working at the Centrum Wiskunde & Informatica (CWI), Amsterdam, on a model of flood risk, with as goal the computation of optimal investments in dike height.